یادداشت‌های جبر خطی

سارا رجب‌زاده
سید پارسا نشایی
مهدی صالح

فهرست مطالب

۱
مقدمه

جبرخطی، شاخه‌ای از ریاضیات است که به بررسی ماتریس‌ها، بردارها، تبدیلات خطی و... می‌پردازد. اکنون عصر داده‌ها آغاز شده است و بنابراین یادگیری زبان ماتریس‌ها و بردارها به یکی از ملزومات مهم برای تحقیق و توسعه بر روی داده‌ها تبدیل شده است.

۱.۱ کاربردها

جبرخطی، تقریبا در تمام زمینه‌های ریاضیات کاربرد دارد. در اینجا، برخی از کاربردهای مهم‌تر آن در علوم کامپیوتر را معرفی می‌کنیم.

رتبه‌بندی صفحات در موتورهای جست‌وجو

در رتبه‌بندی صفحات در موتورهای جست‌وجو، سعی می‌شود که صفحات مرتبط‌تر با پرسمان کاربر به رتبه‌های بالاتری دست پیدا کنند تا کاربر بتواند راحت‌تر به صفحات مورد نظر خود دست یابد. برای پیاده‌سازی این الگوریتم‌ها، از جبرخطی استفاده می‌شود.

تشخیص چهره

یک روش برای تشخیص چهره به صورت خودکار، استفاده از الگوریتم‌ها و متدهایی هم‌چون PCA است که در طراحی آن‌ها، از جبرخطی استفاده می‌شود.

یادگیری ماشین

استفاده از ماتریس‌ها در کاربردهای یادگیری ماشین و یادگیری عمیق که از زیرشاخه‌های هوش مصنوعی هستند، بسیار مهم و فراوان است.

حل مسائل برنامه‌ریزی خطی

در مسائل برنامه‌ریزی خطی به دنبال کمینه یا بیشینه کردن هزینه‌ی لازم برای انجام یک کار (مثلا کمینه کردن طول مسیر از خانه تا دانشگاه) هستیم، که به کمک ابزارهای جبرخطی می‌توان این مسائل را مدل کرد.

واقعیت افزوده

برای قرار دادن اشیای مجازی در فضای حقیقی، نیاز به درک عمیق تبدیلاتی هندسی داریم که می‌توان آن‌ها را به کمک ماتریس‌ها نمایش داد.

تجزیه و تحلیل سیگنال‌ها

ساده‌ترین راه درک تبدیل فوریه (که در تجزیه و تحلیل سیگنال‌ها بسیار مورد استفاده قرار می‌گیرد)، استفاده از جبرخطی است.

تحلیل بسیاری از مسائل در فیزیک، شیمی، اقتصاد و مهندسی، منجر به حل دستگاه معادلات غیرخطی می‌شود که اغلب با یک دستگاه معادلات خطی تقریب زده و آنالیز می‌شود. در اکثر مواقع، ضریب مجهولات دستگاه، اعداد حقیقی هستند، اما گاهی نیز اعداد مختلط هستند. هم‌چنین، حل **معادلات خطی** نیز از جمله مسائل اساسی جبرخطی به حساب می‌آید.

فرض کنید که تعداد مجهول‌ها با تعداد معادله‌ها در یک دستگاه خطی، برابر باشند (یا به عبارتی، n معادله بر حسب n مجهول داشته باشیم). یک مثال از این نوع دستگاه‌های خطی را برای $n = ۲$ در زیر مشاهده می‌کنید:

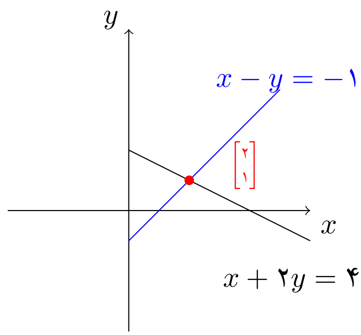
$$\begin{cases} x - y = ۱ \\ x + ۲y = ۴ \end{cases}$$

یک دستگاه معادلات خطی را می‌توان به فرم‌های مختلفی نمایش داد:

۱ ــ فرم ماتریسی: در این فرم، ضرایب مجهول‌ها را به صورت یک ماتریس در نظر می‌گیریم:

$$\begin{bmatrix} ۱ & -۱ \\ ۱ & ۲ \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ۱ \\ ۴ \end{bmatrix}$$

۲ـ تصویر سطری: هر سطر دستگاه، تعبیری هندسی در فضای R^n دارد. با رسم هر سطر در R^n، می‌توان در رابطه با جواب دستگاه قضاوت کرد:



۳ـ تصویر ستونی: می‌توانیم دستگاه را بر اساس ستون‌های ماتریس ضرایبش نمایش دهیم:

$$x \begin{bmatrix} ۱ \\ ۱ \end{bmatrix} + y \begin{bmatrix} -۱ \\ ۲ \end{bmatrix} = \begin{bmatrix} ۱ \\ ۴ \end{bmatrix}$$

به عبارت دیگر، حل دستگاه خطی، معادل با پاسخ دادن به این سوال است که آیا می‌توان بردار $\begin{bmatrix} ۱ \\ ۴ \end{bmatrix}$ را به صورت ترکیب خطی بردارهای $\begin{bmatrix} -۱ \\ ۲ \end{bmatrix}$ و $\begin{bmatrix} ۱ \\ ۱ \end{bmatrix}$ نوشت یا خیر.

فرض کنید $Ax = b$ یک دستگاه شامل n معادله و n مجهول است. در این صورت با استفاده از تصویر ستونی، می‌توان این سوال را مطرح کرد که به ازای چه مجموعه‌ای از بردارهای $b \in R^n$، بردار b را می‌توان به صورت ترکیب خطی ستون‌های ماتریس A نوشت و معادلاً، به ازای چه بردارهای $b \in R^n$، دستگاه $Ax = b$ جواب دارد.

تعبیر تابع خطی دستگاه‌های خطی

فرض کنید $Ax = b$. تابع $T : R^n \to R^n$ با ضابطه‌ی $T(x) = Ax$ را در نظر بگیرید. برد تابع T را با نماد $Im(T)$ نمایش می‌دهند و $Im(T) = \{Ax | x \in R^n\}$. بنابراین $Ax = b$ جواب دارد اگر و تنها اگر $b \in Im(T)$.

تابع T، دو خاصیت مهم دارد: به ازای هر $x, y \in R^n$ و $c \in R$، $T(x + y) = T(x) + T(y)$

و $T(cx) = cT(x)$
هر تابعی مانند $T : R^n \xrightarrow{R^n}$ که دارای دو خاصیت فوق باشد را تابع خطی می‌نامند. مجموعه‌ی همه‌ی تابع‌های خطی از R^n به R^n را با $l(R^n, R^n)$ نمایش می‌دهند. حال با توجه به این که هر بردار مانند V در R^n را می‌توان به صورت ترکیب خطی مجموعه بردارهای $\{e_۱, \cdots, e_n\}$ ــ که در آن، e_iبرداری در R^n است که همه‌ی مؤلفه‌های آن به غیر از مؤلفه‌ی iام که یک است، صفر هستند ــ نوشت، این سوال به ذهن خطور می‌کند که آیا $T(v)$ را نیز می‌توان به صورت ترکیبی خطی از $T(e_۱), \cdots, T(e_n)$ نوشت؟
پاسخ این سوال، «بله» است؛ زیرا اگر فرض کنید $v = c_۱e_۱ + \cdots + c_ne_n$، آن‌گاه داریم:

$$T(v) = c_۱T(e_۱) + \cdots + c_nT(e_n)$$

بنابراین، اگر تابعی خطی روی R^n داشته باشیم، ضابطه‌ی آن به ازای n تا مقدار $T(e_۱), \cdots, T(e_n)$ مشخص می‌شود و به عبارت دیگر، ضابطه‌ی T را با در نظر گرفتن $\{e_۱, \cdots, e_n\}$ را به عنوان پایه (به این معنا که هر بردار در R^n را می‌توان بر حسب اعضای پایه نوشت و هیچ یک از اعضای پایه را نمی‌توان بر حسب سایر اعضای آن نوشت) برای R^n در نظر می‌گیریم. آن‌گاه:

$$T(v) = c_۱T(e_۱) + \cdots + c_nT(e_n) = \begin{bmatrix} T(e_۱) & \cdots & T(e_n) \end{bmatrix} \begin{bmatrix} c_۱ \\ \vdots \\ c_n \end{bmatrix}$$

سوالی طبیعی که به ذهن می‌رسد، این است که اگر پایه‌ی دیگری (مجموعه‌ای دیگر از بردارها با دو ویژگی ذکر شده) برای R^n در نظر بگیریم، ماتریس نمایش T در پایه‌ی جدید، چه ارتباطی با ماتریس نمایش T در پایه‌ی قبلی دارد؟
فرض کنید ماتریس نمایش T در پایه‌ی قدیم را با $[T]_O$ و ماتریس نمایش T در پایه‌ی جدید را با $[T]_N$ نمایش دهیم. نشان خواهیم داد که ماتریس وارون‌پذیر P وجود دارد که $[T]_N = P^{-۱}[T]_OP$، به عبارت دیگر، می‌گوییم دو ماتریس نمایش در پایه‌های مختلف، ماتریس‌هایی **مشابه** هستند.

فضاهای خطی
R^n مجموعه‌ای از بردارهاست که روی آن‌ها جمع برداری و ضرب اسکالر تعریف شده است که

این دو عملگر روی R^n ویژگی‌هایی از جمله جابه‌جایی نسبت به جمع، شرکت‌پذیری، عضو خنثی جمعی، وارون جمعی، عضو خنثی ضربی و توزیع‌پذیری را دارند. به عبارت خودمانی، هر بردار v، $w \in R^n$، $c \in R$، $cv + w \in R^n$. R^n را «فضای خطی» می‌نامند.
حال می‌خواهیم مثال‌های دیگری از مجموعه‌هایی از اشیا ارائه کنیم که ویژگی‌های مذکور را داشته باشند.

مثال: فرض کنید

$$P_n(x) = \{a_n x^n + \cdots + a_۱ x + a_۰ | a_i \in R\, \forall ۱ \leq i \leq n\}$$

به عبارت دیگر، در این مثال $P(x)$ مجموعه‌ی همه‌ی چند جمله‌ای‌های با درجه‌ی حداکثر n است؛ که بر اساس آن‌چه گفته شد، مشخصا یک فضای خطی است.
مثال: تابع مشتق $\frac{d}{dx} : P_۳(x) \to P_۳(x)$ یک تابع خطی است. هر چندجمله‌ای حداکثر از درجه‌ی ۳ را می‌توان به صورت ترکیب خطی $E = \{x^۳, x^۲, x, ۱\}$ نوشت؛ بنابراین، نمایش تابع خطی $\frac{d}{dx}$ برابر است با

$$\begin{bmatrix} ۰ & ۰ & ۰ & ۰ \\ ۳ & ۰ & ۰ & ۰ \\ ۰ & ۲ & ۰ & ۰ \\ ۰ & ۰ & ۱ & ۰ \end{bmatrix}$$

به طور کلی، فرض کنید V یک فضای خطی باشد. مجموعه‌ی همه‌ی توابع خطی از V به V را با نماد $l(V, V)$ نمایش می‌دهیم. به صورت کلی‌تر، اگر V و W فضاهای خطی باشند، مجموعه‌ی همه‌ی توابع خطی از V به W را با $l(V, W)$ نشان می‌دهیم.
حال فرض کنید $U \subseteq W$ و $T : V \to V$ تبدیل خطی باشد. اثر تبدیل خطی T روی U (که U زیرفضایی از V باشد، به این معنا که ویژگی‌های فضای برداری را داشته باشد) را بررسی می‌کنیم. اگر به ازای هر $u \in U$، $T(u)$ عضوی از U باقی بماند می گوییم تحدید T به U ناوردا است.
فرض کنید که v برداری ناصفر از V باشد و U زیرفضایی از V باشد که هر عضو آن، مضربی از بردار v است؛ یعنی $U = \{cv | c \in R\}$ که مشخصا زیرفضایی از V خواهد شد. حال سوال این است که تحت چه شرایطی، $T(U) \subseteq U$.
پاسخ به این سوال، منجر به تعریف رده‌ی مهمی از بردارها در V می‌گردد. اگر $v \in V$ برداری

باشد که $T(v) = \lambda v$، آنگاه v را **بردار ویژه** و λ را **مقدار ویژه** نگاشت A می‌نامند. به طور طبیعی، این مفهوم را می‌توان برای ماتریس‌ها نیز تعریف کرد؛ فرض کنید A ماتریسی $n * n$ باشد و $T(x) = Ax$؛ بنابراین، $v \neq 0$ بردار ویژه‌ی A است، اگر و تنها اگر $T(v) = \lambda v$؛ در فرم ماتریسی نیز λ را مقدار ویژه‌ی A می‌نامند.

اکنون خوب است به این پرسش توجه نماییم که آیا می‌توان فضای V را به‌گونه‌ای برحسب زیرفضاهایی که تحت نگاشت A ناوردا هستند، «تجزیه» کرد؟ مشخصاً، آیا می‌توان این کار را برحسب زیرفضاهایی که با بردارهای ویژه از یک‌دیگر متمایز می‌شوند، انجام داد؟

هر چه در تجزیه‌ی فضا برحسب زیرفضاهای ناوردا موفق‌تر باشیم، به نمایش ساده‌تری از ماتریس متناظر با نگاشت خطی T دست خواهیم یافت.

به طور مثال، فرض کنید $T : V \to V$ یک تبدیل خطی باشد و $B = \{v_1, \cdots, v_n\}$ پایه‌ای (به همان معنا که گفته شد) باشد، به طوری که $T(v_1) = \lambda_1 v_1$؛ در این صورت نمایش تبدیل خطی T در پایه‌ی B

$$\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

یک ماتریس قطری خواهد شد!

محاسبه‌ی فاصله روی فضاهای خطی

در R^n فاصله‌ی هر دو بردار $w, v \in R^n$، قابل محاسبه است. به طور خاص، فاصله‌ی هر بردار از مبدأ که طول بردار نامیده می‌شود، قابل محاسبه است (حداقل با دانشی که تا کنون داریم و طول اقلیدسی) و می‌توان زاویه‌ی بین دو بردار را نیز محاسبه کرد. توسیع این مفاهیم به هر فضای برداری مانند V، با تعریف ضرب داخلی انجام می‌شود. (اسلایدهای ۲۳ تا ۲۵)

۲

حذف گاوسی

۱.۲ حل دستگاه‌ها

برای حل دستگاه‌های چند معادله و چند مجهول، می‌توان از روش حذف گاوسی[1] استفاده کرد.

در این روش، ابتدا ضرایب متغیرهای مختلف را به صورت یک **ماتریس** نوشته و بردار پاسخ معادلات را نیز برای راحتی به سمت راست ماتریس، ملحق [2] می‌کنیم. در مرحله‌ی اول، سطر اول، در مرحله‌ی دوم، سطر دوم و... را به عنوان سطرهای ثابت در آن مرحله در نظر می‌گیریم و در هر مرحله اطمینان حاصل می‌کنیم که درایه‌ی روی قطر اصلی در سطر ثابت، صفر **نباشد**. اگر آن درایه، صفر بود، در صورت امکان، آن سطر را با یکی از سطرهای زیر آن که همان درایه‌اش صفر نیست، جابه‌جا می‌کنیم و ادامه می‌دهیم. سپس، ضریبی از سطر ثابت را از باقی سطرها کم می‌کنیم تا درایه‌های زیر قطر اصلی آن‌ها برابر با صفر شود و در نهایت، ماتریس تبدیل به یک ماتریس **بالامثلثی** شود.

مثال: دستگاه معادلات زیر را در نظر بگیرید:

$$\begin{cases} ۲u + v + w &= ۵ \\ ۴u - ۶v &= -۲ \\ -۲u + ۷v + ۲w &= ۹ \end{cases}$$

برای حل دستگاه بالاِ، همان‌گونه که پیش‌تر ذکر شد، ابتدا ماتریس ضرایب را تشکیل می‌دهیم؛

1) Gaussian Elimination 2) augment

هم‌چنین، بردار $b = \begin{bmatrix} ۵ \\ -۲ \\ ۹ \end{bmatrix}$ را نیز در کنار ماتریس ضرایب، می‌افزاییم.

$$A = \begin{bmatrix} ۲ & ۱ & ۱ & ۵ \\ ۴ & -۶ & ۰ & -۲ \\ -۲ & ۷ & ۲ & ۹ \end{bmatrix}$$

در معادله‌ی اول، ضریب متغیر u عدد ۲ است که با استفاده از آن ضرایب متغیر u در سایر معادلات صفر خواهد شد. ابتدا در مرحله‌ی اول، دو برابر سطر اول را از سطر دوم کم می‌کنیم و −۱ برابر سطر اول را از سطر سوم کم می‌کنیم:

$$\begin{bmatrix} ۲ & ۱ & ۱ & ۵ \\ ۰ & -۸ & -۲ & -۱۲ \\ ۰ & ۸ & ۳ & ۱۴ \end{bmatrix}$$

در مرحله‌ی بعدی، ضریب v در معادله‌ی دوم برابر با −۸ است که با استفاده از آن، باقی ضرایب v در معادلات پایین‌تر را صفر می‌کنیم. برای این کار −۱ برابر سطر دوم را از سطر سوم کم می‌کنیم:

$$\begin{bmatrix} ۲ & ۱ & ۱ & ۵ \\ ۰ & -۸ & -۲ & -۱۲ \\ ۰ & ۰ & ۱ & ۲ \end{bmatrix}$$

حال، از معادله‌ی سوم به سمت بالا حرکت می‌کنیم و پاسخ‌ها را به دست می‌آوریم:

$$w = ۲$$

$$v = ۱$$

$$u = ۱$$

نکته: بنا به تعریف، درایه‌های محوری، نمی‌توانند صفر باشند.

پرسش:
فرایند فوق، تحت چه شرایطی به شکست می‌انجامد؟
پاسخ:
فرض کنید که دستگاه n معادله و n مجهول زیر داده شده است:

$$\begin{cases} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & \vdots & = & \vdots \\ a_{n1}x_1 & + & \cdots & + & a_{nn}x_n & = & b_n \end{cases}$$

۱. در مرحله‌ی اول، اگر $a_{11} \neq 0$ فرایند صفر کردن ضرایب x_1 در سایر معادلات انجام می‌شود، ولی اگر $a_{11} = 0$ آنگاه به ضرایب x_1 در سایر معادلات نگاه می‌کنیم و معادله‌ای که ضریب x_1 در آن ناصفر است را به جای معادله‌ی اول قرار می‌دهیم (جای دو معادله را عوض می‌کنیم). توجه کنید که چون فرض کرده‌ایم دستگاه n مجهول دارد، حداقل یکی از ضرایب x_i ناصفر است.

۲. در مرحله ی i ام $(1 \leq i \leq n-1)$ اگر ضریب x_i صفر باشد به ضریب x_i در معادلات بعدی نگاه می‌کنیم؛ دو حالت امکان‌پذیر است:
الف) یا ضرایب x_i در همه‌ی حالات بعدی صفر است که در این صورت، به مرحله ی $i+1$ می رویم (اگر $i+1 \leq n-1$) یا این‌که یکی از ضرایب ناصفر باشد، این دو معادله را جابه‌جا می‌کنیم و ضریب ناصفر، درایه‌ی محوری i ام خواهد بود و سپس، فرایند صفر کردن ضرایب x_i در معادلات بعدی را انجام می‌دهیم.

مثال:

$$\begin{cases} u+v+w & = & b_1 \\ 2u+2v+5w & = & b_2 \\ 4u+6v+8w & = & b_3 \end{cases}$$

$$\begin{bmatrix} 1 & 1 & 1 & b_1 \\ 2 & 2 & 5 & b_2 \\ 4 & 6 & 8 & b_3 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} ۱ & ۱ & ۱ & b_۱ \\ ۰ & ۰ & ۳ & b_۲ - ۲b_۱ \\ ۰ & ۲ & ۴ & b_۳ - ۴b_۱ \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} ۱ & ۱ & ۱ & b_۱ \\ ۰ & ۲ & ۴ & b_۳ - ۴b_۱ \\ ۰ & ۰ & ۳ & b_۲ - ۲b_۱ \end{bmatrix}$$

حال u و v و w قابل محاسبه هستند و دستگاه، جواب یکتا دارد.

مثال:

$$\begin{cases} u + v + w & = & b_۱ \\ ۲u + ۲v + ۵w & = & b_۲ \\ ۴u + ۴v + ۸w & = & b_۳ \end{cases}$$

$$\begin{bmatrix} ۱ & ۱ & ۱ & b_۱ \\ ۲ & ۲ & ۵ & b_۲ \\ ۴ & ۴ & ۸ & b_۳ \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} ۱ & ۱ & ۱ & b_۱ \\ ۰ & ۰ & ۳ & b_۲ - ۲b_۱ \\ ۰ & ۰ & ۴ & b_۳ - ۴b_۱ \end{bmatrix}$$

از معادله‌های دوم و سوم داریم:

$$\begin{cases} w = \dfrac{b_۳ - ۴b_۱}{۴} \\ \\ w = \dfrac{b_۲ - ۲b_۱}{۳} \end{cases}$$

دستگاه با جابه‌جایی سطر، قابل اصلاح نیست. اگر دو جواب به دست آمده برای w با یک‌دیگر برابر باشند، از معادله‌ی اول داریم:

$$u = b_۱ - w - v$$

در نتیجه، دستگاه **بی‌شمار** جواب دارد؛ اما اگر آن دو جواب با یک‌دیگر برابر نباشند، دستگاه جواب **ندارد**.

محاسبه هزینه‌ی حذف گاوسی:
فرض کنید ماتریس زیر از یک دستگاه n معادله و n مجهول به دست آمده باشد:

$$\begin{bmatrix} a_{۱۱} & \cdots & a_{۱n} & b_۱ \\ \vdots & & & \vdots \\ a_{n۱} & \cdots & a_{nn} & b_n \end{bmatrix}$$

هم‌چنین، فرض کنید هر تقسیم، ضرب و یا تفریق را را یک عمل حساب کنیم؛ در این صورت:

- ستون اول: با استفاده از $a_{۱۱} \neq ۰$ همه‌ی درایه‌های این ستون، به غیر از $a_{۱۱}$ باید صفر شود، پس همه درایه‌های ماتریس، به غیر از درایه‌های سطر اول، دستخوش تغییر قرار می‌گیرند. پس در مرحله‌ی اول حذف $n^۲ - n$ عمل انجام می‌شود.

- ستون دوم: با استفاده از درایه‌ی دوم ستون دوم (در صورت ناصفر بودن) درایه‌های زیر سطر دوم در ستون دوم صفر می‌شوند. پس $(n - ۱)^۲ - (n - ۱)$ عمل انجام می‌شود.

از طرفی، این اعمال روی ستون $\begin{bmatrix} b_۱ \\ \vdots \\ b_n \end{bmatrix}$ نیز انجام می‌شوند؛ در مرحله‌ی اول $n - ۱$ عمل، در مرحله‌ی دوم $n - ۲$ عمل و...

بنابراین، حداکثر اعمال مورد نیاز برای تشکیل ماتریس U (ماتریس بالا مثلثی نتیجه شده از این اعمال) برابر است با

$$((n^2-n)+((n-1)^2-(n-1))+\cdots+1)+((n-1)+(n-2)+\cdots+1)=\frac{n^3-n}{3}+\frac{n(n-1)}{2}.$$

توجه شود که اگر در مرحله‌ای، جابه‌جایی سطری نیز لازم بود، آن را انجام می‌دهیم.

هم‌چنین، برای محاسبه‌ی جواب آخر، داریم:

$$Ux=\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

برای محاسبه‌ی x_n یک عمل، برای محاسبه‌ی x_{n-1} دو عمل، ... و برای محاسبه‌ی x_1 به n عمل نیاز داریم؛ پس:

$$n+(n-1)+(n-2)+\cdots+1=\frac{n(n+1)}{2}$$

پس مجموع هزینه‌ی محاسبه، برابر است با:

$$\frac{n^3-n}{3}+\frac{n(n-1)}{2}+\frac{n(n+1)}{2}=\frac{n^3+3n^2-n}{3}\approx\frac{1}{3}n^3$$

۲.۲ تجزیه‌ی LU

در انجام عملیات حذف گاوسی، به یک ماتریس بالامثلثی می‌رسیم؛ از طرفی، یک سری عملیات سطری روی ماتریس اولیه انجام داده‌ایم که هر یک از این عملیات سطری، خود با یک ماتریس مدل می‌شوند؛ در نتیجه اگر ماتریس‌های مربوط به عملیات سطری را E_1 تا E_n بنامیم، داریم:

$$E_1E_2\ldots E_nA=U\rightarrow A=(E_1E_2\ldots E_n)^{-1}U$$

که اگر در آن، تعریف کنیم $L=(E_1E_2\ldots E_n)^{-1}$ آن‌گاه $A=LU$ و یک تجزیه برای A به صورت ضرب دو ماتریس یافته‌ایم.

برای حل دستگاه‌های خطی، می‌توان ابتدا ماتریس ضرایب را به L که پایین مثلثی است و U که بالا مثلثی است، تجزیه، و سپس معادلات خطی $Lc = b$ و $Ux = c$ را حل کنیم. دلیل پایین‌مثلثی بودن L آن است که از وارون ضرب E_i ها که هر یک پایین‌مثلثی هستند، به دست آمده است و ثابت می‌شود که ضرب و وارون ماتریس‌های پایین مثلثی، ماتریسی پایین مثلثی است.

تمرین: ماتریس زیر را به L و U تجزیه کنید.

$$A = \begin{bmatrix} ۱ & ۲ & ۳ \\ ۱ & ۴ & ۶ \\ ۴ & ۵ & ۶ \end{bmatrix}$$

پاسخ:

$$L = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۱ & ۱ & ۰ \\ ۴ & -۱/۵ & ۱ \end{bmatrix}$$

$$U = \begin{bmatrix} ۱ & ۲ & ۳ \\ ۰ & ۲ & ۳ \\ ۰ & ۰ & -۱/۵ \end{bmatrix}$$

$$A = LU$$

در صورت لزوم، می‌توانیم یک ماتریس را به مولفه‌های L و D و U نیز تجزیه کنیم، به طوری‌که مولفه‌های قطر اصلی U روی قطر اصلی D که یک ماتریس قطری است، ظاهر شده و تمامی سطرهای U بر درایه‌ی محوری‌اش تقسیم می‌شود تا درایه‌های محوری ماتریس U ی حاصل، همگی یک شوند. در تمرین فوق،

$$L = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۱ & ۱ & ۰ \\ ۴ & -۱/۵ & ۱ \end{bmatrix}$$

$$D = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۲ & ۰ \\ ۰ & ۰ & -۱/۵ \end{bmatrix}$$

$$U = \begin{bmatrix} ۱ & ۲ & ۳ \\ ۰ & ۱ & ۱/۵ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

$$A = LDU$$

ممکن است در حذف گاوسی، نیاز به جابه‌جایی سطرها نیز داشته باشیم، که در تجزیه‌ی LU به صورت $PA = LU$ ظاهر می‌شود که در آن P یک ماتریس «جایگشت» است که از جابه‌جایی سطرهای ماتریس همانی به دست آمده است.

مثال:

تحت چه شرایطی، حاصل ضرب زیر، وارون‌پذیر است؟

$$A = \begin{bmatrix} ۱ & ۰ & ۰ \\ -۱ & ۱ & ۰ \\ ۰ & -۱ & ۱ \end{bmatrix} \begin{bmatrix} d_۱ & & \\ & d_۲ & \\ & & d_۳ \end{bmatrix} \begin{bmatrix} ۱ & -۱ & ۰ \\ ۰ & ۱ & -۱ \\ ۰ & ۰ & ۱ \end{bmatrix}$$

پاسخ:

مشخص است که تجزیه‌ی LDU برای ماتریس A داده شده است، پس برای وارون‌پذیر بودن، باید درایه‌های محوری که در قطر اصلی D واقع‌اند (یعنی d_i ها) ناصفر باشند.

مثال:

چه c ای باعث می‌شود در درایه‌ی محوری دوم و سوم ماتریس زیر، صفر ایجاد شود؟ (پرسش مطرح شده در اسلاید ۱۷ درس)

$$\begin{bmatrix} ۱ & c & ۰ \\ ۲ & ۴ & ۱ \\ ۳ & ۵ & ۱ \end{bmatrix}$$

پاسخ:

اگر سطر دوم بخواهد درایه‌ی محوری صفر پیدا کند، باید حتما $c = ۲$ باشد تا وقتی دو برابر آن از ۴ کم می‌شود، صفر شود. اگر این‌طور نباشد، ماتریس پس از حذف زیر محور ستون اول به شکل زیر در می‌آید:

$$\begin{bmatrix} ۱ & c & ۰ \\ ۰ & ۴ - ۲c & ۱ \\ ۰ & ۵ - ۳c & ۱ \end{bmatrix}$$

حال اگر $۵ - ۳c = ۴ - ۲c$ آنگاه با حذف زیر محور ستون دوم، درایه‌ی محوری سطر سوم، صفر می‌شود (زیرا یک منهای یک برابر صفر است)؛ در نتیجه به دست می‌آوریم

$$c = ۱$$

مثال:

ماتریس زیر را به L و U تجزیه کنید. در چه شرایطی، ماتریس چهار محور دارد؟

$$A = \begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix}.$$

پاسخ:

با حذف گاوسی، به ماتریس‌های زیر می‌رسیم:

$$L = \begin{bmatrix} ۱ & ۰ & ۰ & ۰ \\ ۱ & ۱ & ۰ & ۰ \\ ۱ & ۱ & ۱ & ۰ \\ ۱ & ۱ & ۱ & ۱ \end{bmatrix}.$$

$$U = \begin{bmatrix} a & a & a & a \\ ۰ & b-a & b-a & b-a \\ ۰ & ۰ & c-b & c-b \\ ۰ & ۰ & ۰ & d-c \end{bmatrix}.$$

حال، باید درایه‌های قطری U ناصفر باشند تا ماتریس چهار محور داشته باشد، یعنی

$$a \neq ۰, a \neq b, b \neq c, c \neq d$$

مثال:

ماتریس زیر را به L و U تجزیه کنید. در چه شرایطی، ماتریس چهار محور دارد؟

$$A = \begin{bmatrix} a & r & r & r \\ a & b & s & s \\ a & b & c & t \\ a & b & c & d \end{bmatrix}.$$

پاسخ:

با حذف گاوسی، به ماتریس‌های زیر می‌رسیم:

$$L = \begin{bmatrix} ۱ & ۰ & ۰ & ۰ \\ ۱ & ۱ & ۰ & ۰ \\ ۱ & ۱ & ۱ & ۰ \\ ۱ & ۱ & ۱ & ۱ \end{bmatrix}.$$

$$U = \begin{bmatrix} a & r & r & r \\ 0 & b-r & s-r & s-r \\ 0 & 0 & c-s & t-s \\ 0 & 0 & 0 & d-t \end{bmatrix}.$$

حال، باید درایه‌های قطری U ناصفر باشند تا ماتریس چهار محور داشته باشد، یعنی

$$a \neq 0, b \neq r, c \neq s, d \neq t$$

۳.۲ ماتریس وارون

فرض کنید مجموعه‌ی همه‌ی ماتریس‌های با m سطر و n ستون (ماتریس‌های $m \times n$) با درایه‌های حقیقی را با نماد $M_{mn}(R)$ و نیز مجموعه‌ی همه‌ی ماتریس‌های با n سطر و n ستون را با $M_n(R)$ نمایش دهیم.

تعریف: اگر به ازای ماتریس $A \in M_n(R)$ ، ماتریس $B \in M_n(R)$ وجود داشته باشد به طوری که $AB = BA = I$، آنگاه می‌گوییم A **وارون‌پذیر** است و وارون آن B خواهد بود.

نکته: لزوما، همه‌ی ماتریس‌ها وارون ندارند.

گزاره: اگر $A, B, C \in M_n(R)$ به طوری که $CA = I$ و $AB = I$، آنگاه $B = C$.

برهان: اگر C را در دو طرف تساوی $AB = I$ ضرب کنیم، داریم $CAB = C$. مطابق فرض گزاره، می‌دانیم که $CA = I$ است، بنابراین نتیجه می‌گیریم $IB = C$ و در نتیجه $B = C$.

نتیجه: اگر ماتریس $A \in M_n$ وارون داشته باشد، وارون آن یکتا است و آن را با نماد A^{-1} نمایش می‌دهیم.

نکته: اگر A وارون‌پذیر باشد، دستگاه خطی $Ax = b$ دارای جواب یکتای $x = A^{-1}b$ است.

گزاره: اگر $A_1, \cdots , A_k$ ماتریس‌هایی وارون‌پذیر باشند، حاصل‌ضرب آن‌ها نیز وارون‌پذیر خواهد بود.

برهان:

$$(A_1 A_2 \cdots A_k)(A_k^{-1} \cdots A_2^{-1} A_1^{-1}) = I$$

$$(A_k^{-1} \cdots A_2^{-1} A_1^{-1})(A_1 A_2 \cdots A_k) = I$$

در نتیجه $A_1 \cdots A_k$ وارون‌پذیر است و وارون آن برابر است با:

$$(A_1 \cdots A_k)^{-1} = A_k^{-1} \cdots A_2^{-1} A_1^{-1}$$

۴.۲ محاسبه‌ی A^{-1}: روش گاوس ـ ژردان

[1] فرض کنید ماتریسی $n \times n$ به نام A داریم. ابتدا ماتریس همانی $n \times n$ را در کنار ماتریس A به گونه‌ای قرار می‌دهیم که به ماتریسی $n \times 2n$ برسیم و سپس هر سطر را به گونه‌ای تغییر می‌دهیم که نیمه‌ی سمت چپ ماتریس به ماتریس همانی تبدیل شود.

مثال: می‌خواهیم وارون ماتریس هیلبرت [2] 3×3 را به دست آوریم.

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & 0 & 0 & 1 \end{bmatrix}$$

حل:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ 0 & \frac{1}{12} & \frac{1}{12} & -\frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{12} & \frac{4}{45} & -\frac{1}{3} & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ 0 & 1 & 1 & -6 & 12 & 0 \\ 0 & 0 & \frac{1}{180} & \frac{1}{6} & -1 & 1 \end{bmatrix}$$

[1] Gauss-Jordan
[2] Hilbert matrix

$$\left[\begin{array}{ccc|ccc} ۱ & ۰ & -\frac{۱}{۶} & ۴ & -۶ & ۰ \\ ۰ & ۱ & ۱ & -۶ & ۱۲ & ۰ \\ ۰ & ۰ & ۱ & ۳۰ & -۱۸۰ & ۱۸۰ \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} ۱ & ۰ & ۰ & ۹ & -۳۶ & ۳۰ \\ ۰ & ۱ & ۰ & -۳۶ & ۱۹۲ & -۱۸۰ \\ ۰ & ۰ & ۱ & ۳۰ & -۱۸۰ & ۱۸۰ \end{array}\right]$$

$$H^{-۱} = \begin{bmatrix} ۹ & -۳۶ & ۳۰ \\ -۳۶ & ۱۹۲ & -۱۸۰ \\ ۳۰ & -۱۸۰ & ۱۸۰ \end{bmatrix}$$

ترانهاده‌ی ماتریس

ترانهاده‌ی یک ماتریس $A_{n\times m}$، ماتریسی $m \times n$ است که آن را با نماد A^T نمایش می‌دهیم، و به ازای هر i و j داریم $A^T_{ij} = A_{ji}$. به عبارتی، ترانهاده‌ی ماتریس A، ماتریسی است که در آن، جای سطرها و ستون‌های A عوض شده است.

مثال:

$$A = \begin{bmatrix} ۲ & -۳ & ۵ \\ ۰ & ۹ & ۱۰ \end{bmatrix} \qquad A^T = \begin{bmatrix} ۲ & ۰ \\ -۳ & ۹ \\ ۵ & ۱۰ \end{bmatrix}$$

در حالت کلی داریم:

$$A = \begin{bmatrix} a_{۱۱} & a_{۱۲} & \dots & a_{۱n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m۱} & a_{m۲} & \dots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{۱۱} & \dots & a_{m۱} \\ a_{۱۲} & \dots & a_{m۲} \\ \vdots & \vdots & \vdots \\ a_{۱n} & \dots & a_{mn} \end{bmatrix}$$

چند نکته‌ی مهم:

۱. همواره داریم $(AB)^T = B^T A^T$

۲. همواره داریم $(A^T)^{-۱} = (A^{-۱})^T$

۳. ماتریسی که با ترانهاده‌اش برابر باشد، ماتریس **متقارن** نام دارد.

۴. ماتریس‌های متقارن، همواره مربعی هستند، زیرا در غیر این صورت، ماتریس و ترانهاده‌اش، ابعاد متفاوتی دارند و در نتیجه برابر نیستند و بنابراین، ماتریس متقارن نخواهد بود.

۵. لزومی ندارد که ماتریس‌های متقارن، وارون‌پذیر نیز باشند.

گزاره: ماتریس A متقارن است، اگر و تنها اگر ماتریس $A^{-۱}$ متقارن باشد.

برهان: می‌دانیم I متقارن است، پس:

$$I = I^T \to AA^{-۱} = (AA^{-۱})^T = (A^{-۱})^T A^T$$

$$\to A^{-۱}A = (A^{-۱})^T A^T$$

چون A متقارن است، $A = A^T$، در نتیجه می‌توانیم در رابطه‌ی فوق، به جای ترانهاده‌ی A خود A را قرار دهیم:

$$A^{-۱}A = (A^{-۱})^T A \to A^{-۱}AA^{-۱} = (A^{-۱})^T A^T A^{-۱}$$

و چون $AA^{-۱} = I$، نتیجه می‌گیریم $A^{-۱} = (A^{-۱})^T$ و در نتیجه وارون A متقارن است. حکم دو شرطی است، زیرا می‌توانیم در اثبات فوق، همه‌جا به جای A وارون آن را قرار دهیم تا دو شرطی بودن حکم، ثابت شود.

کاربرد ماتریس‌ها در حل معادلات دیفرانسیل

فرض کنید یک معادله‌ی دیفرانسیل با شرایط مرزی (مقادیر تابع، و نه مشتق آن، در دو سر بازه) داده شده است. می‌دانیم اگر مقادیر مشتق‌های تابع در نقطه‌ی اولیه مشخص باشند، می‌توانیم به کمک روش‌های آموخته‌شده در درس معادلات دیفرانسیل، این معادلات را حل کنیم، اما در حل معادلات با شرایط مرزی، استفاده از جبرخطی راه‌گشا است. با توجه به ماهیت گسسته‌ی ماتریس‌ها، مقدار تابع را در n نقطه با فاصله‌ی یکسان از هم در نظر می‌گیریم و مقادیر تقریبی برای جواب اصلی را در این نقاط می‌یابیم. می‌توان میزان اختلاف و تغییرات تابع را به دو روش «رو به جلو» و «رو به عقب» مدل کرد:

$$\frac{\Delta u}{\Delta x} = \frac{u(x+h) - u(x)}{h} \text{ or } \frac{u(x) - u(x-h)}{h}$$

$$\begin{aligned}\to \frac{d^۲u}{dx^۲} \approx \frac{\Delta^۲ u}{\Delta x^۲} = \frac{\Delta}{\Delta x}\Big(\frac{\Delta u}{\Delta x}\Big) &= \frac{\frac{u(x+h)-u(x)}{h} - \frac{u(x)-u(x-h)}{h}}{h} \\ &= \frac{u(x+h) - ۲u(x) + u(x-h)}{h^۲}.\end{aligned}$$

با در نظر گرفتن یک h ثابت بین ۰ و ۱ و بخش‌بندی بازهٔ $(0, 1)$ به فواصل به‌طول h، می‌توانیم معادلهٔ فوق را برای هر کدام از مقادیر $u_i = u(ih)$ بازنویسی نماییم. در ان صورت، به دستگاهی از معادلات خطی خواهیم رسید که برای $h = 0/2$ چنین شکلی دارد:

$$\begin{cases} 2u_1 - 1u_2 + = h^2 f(h) \\ -1u_1 + 2u_2 - 1u_3 = h^2 f(2h) \\ -2u_2 + 2u_3 - 1u_4 = h^2 f(3h) \\ -1u_3 + 2u_4 - 1u_5 = h^2 f(4h) \\ -1u_4 + 2u_5 = h^2 f(5h) \end{cases}$$

و فرم ماتریسی آن به این صورت است:

$$\underbrace{\begin{bmatrix} 2 & 1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}}_{A} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} = h^2 \begin{bmatrix} f(h) \\ f(2h) \\ f(3h) \\ f(4h) \\ f(5h) \end{bmatrix}$$

.

ماتریسی که در این روش استفاده می‌شود، دارای چند ویژگی مهم است:

۱. این ماتریس، شامل یک قطر اصلی و دو قطر در دو طرف آن است که می‌توانند مقادیر ناصفر داشته باشند؛ سایر درایه‌های ماتریس، صفر هستند.

۲. ماتریسِ استفاده شده، متقارن است.

۳. تمامی درایه‌های محوری این ماتریس، مثبت هستند؛ به عبارتی، ماتریس مورد استفاده، «معین مثبت» است.

مثال: معادله‌ی دیفرانسیل $-\frac{d^2u}{dx^2} = x + 1$ داده شده‌است. در نظر بگیرید $h = 0/25$. داریم

$$\begin{cases} 2u_1 - u_2 = \frac{1}{16} \times \frac{5}{4} = \frac{5}{64} \\ -u_1 + 2u_2 - u_3 = \frac{1}{16} \times \frac{6}{4} = \frac{6}{64} \\ -u_2 + 2u_3 = \frac{1}{16} \times \frac{7}{4} = \frac{7}{64} \end{cases}$$

.

که فرم ماتریسی این دستگاه معادلات به این شکل است:

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \frac{1}{64} \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$$

۵.۲ خطای گرد کردن

در کامپیوتر تعداد ثابتی رقم نگهداری می‌شود و بنابراین اعداد معمولاً گرد شده هستند که این گرد کردن موجب خطا خواهد شد. برای مثال، فرض کنید کامپیوتری داریم که توانایی نگه داشتن ارقام تا ۳ رقم اعشار را دارد. می‌خواهیم دو عدد اعشاری زیر را با هم جمع کنیم:

$$0/456 + 0/00123 \to 0/457$$

همان‌طور که مشاهده می‌کنید دو رقم پایانی حاصل‌جمع را از دست دادیم و به همین خاطر خطا داریم. توجه کنید که این خطا در بعضی از محاسبات بسیار زیاد شده و قابل قبول نخواهد بود.
حال می‌خواهیم میزان این خطا را در محاسبه‌ی $Ax = b$ در نظر بگیریم. دو ماتریس را در نظر می‌گیریم که یکی تقریبا وارون‌ناپذیر بوده و دیگری وارون‌پذیر است:

$$A = \underbrace{\begin{bmatrix} 1/0 & 1/0 \\ 1/0 & 1/0001 \end{bmatrix}}_{\text{نزدیک به وارون‌ناپذیری}} \qquad B = \underbrace{\begin{bmatrix} 0/0001 & 1/0 \\ 1/0 & 1/0 \end{bmatrix}}_{\text{دور از وارون‌ناپذیری}}$$

اگر بردارهای b را به صورت زیر نزدیک به هم در نظر بگیریم داریم:

$$\begin{array}{rcrcl} u & + & v & = & ۲ \\ u & + & ۱/۰۰۰۱v & = & ۲/۰۰۰۰ \end{array} \qquad \begin{array}{rcrcl} u & + & v & = & ۲ \\ u & + & ۱/۰۰۰۱v & = & ۲/۰۰۰۱ \end{array}$$

بنابراین داریم:

$$u = ۲ \quad v = ۰ \quad , \qquad u = ۱ \quad v = ۱$$

همان‌طور که مشاهده می‌کنید تغییر در پنجمین رقم b باعث تغییر در اولین رقم پاسخ شد.
باید توجه داشت که ماتریس به‌ظاهر بدون مشکلی مانند B نیز می‌تواند با به کار گیری یک الگوریتم نامناسب، رفتاری مشابه A از خود نشان دهد. برای مثال، فرض کنید ماتریس B را به شکل بالا داشته باشیم و همچنین $b = \begin{bmatrix} ۱ \\ ۲ \end{bmatrix}$ باشد.
اگر فرض کنیم درایه‌ی اول (۰/۰۰۰۱) اولین درایه‌ی محوری باشد، با محاسبه‌ای ساده داریم:

$$\begin{bmatrix} ۰/۰۰۰۱ & ۱/۰ & ۱ \\ ۱/۰ & ۱/۰ & ۲ \end{bmatrix} \rightarrow \begin{bmatrix} ۰/۰۰۰۱ & ۱/۰ & ۱ \\ ۰ & -۹۹۹۹ & -۹۹۹۸ \end{bmatrix}$$

$$v = \frac{۹۹۹۸}{۹۹۹۹} \simeq \begin{cases} ۰/۹۹۹۹ \Rightarrow u = ۱ \\ ۱ \quad\ \Rightarrow u = ۰ \end{cases}$$

عبارت بالا به این معناست که عدد به‌دست‌آمده برای v تقریباً برابر با ۰/۹۹۹۹ است، که اگر آن را همان ۰/۹۹۹۹ در نظر بگیریم، $u = ۱$ شده و اگر آن را گرد کرده و برابر با ۱ در نظر بگیریم، $u = ۰$ خواهد شد.
همان‌طور که مشاهده می‌شود، با یک الگوریتم نامناسب، به دو پاسخ کاملاً متفاوت رسیدیم. دلیل این موضوع، انتخاب درایه‌ی اول به عنوان اولین درایه‌ی محوری بود؛ این درایه بسیار نزدیک به صفر است و به همین دلیل، موجب بی‌ثباتی می‌شود. برای حل این مشکل، می‌توان از جابه‌جایی سطرها [۱] استفاده کرد.

[1] Row Exchange

۳

فضاهای خطی

۱.۳ فضاهای خطی

فضای R^n از همه‌ی بردارهای ستونی n مؤلفه‌ای حقیقی تشکیل شده است که در آن، حاصل‌جمع هر دو بردار و حاصل‌ضرب آن‌ها در اسکالر، کماکان عضوی از R^n است؛ به عبارت دیگر، هر ترکیب خطی از بردارها، در فضا قرار داشته باشد. در این شرایط، R^n را یک فضای خطی می‌نامیم.
علاوه بر فضاهای با بعد محدود، فضای بی‌نهایت بعدی R^∞ نیز وجود دارد که بردارهای درون آن به صورت $x = (x_۱, x_۲, \cdots)$ هستند. هر بردار درون آن، بی‌نهایت مؤلفه دارد و به ازای هر دو بردار x و y و هر اسکالر مانند c، بردارهای $x + y$ و cx نیز درون فضا هستند، یعنی:

$$x, y \in R^\infty, c \in R \qquad \Rightarrow \qquad x + y \in R^\infty, cx \in R^\infty$$

پس R^∞ هم یک فضای خطی است. برای مثال، فضای توابع را در نظر بگیرید و فرض کنید $V = \{f : [۰, ۱] \to R\}$. در این صورت، اگر در نظر بگیریم $g, f \in V$ و $c \in R$، می‌توان به راحتی دید که $f + g \in R$ و $cf \in R$.
فضای چندجمله‌ای‌ها از درجه‌ی حداکثر n نیز خواص مشابهی دارد.
تعریف فضای خطی: مجموعه‌ای مانند V به همراه جمع و ضرب اسکالر روی V، $(V, +, \cdot)$، فضای خطی است، اگر و تنها اگر ویژگی‌های زیر درست باشند:

۱. جابه‌جایی: به ازای هر u و v $\qquad u + v = v + u$

۲. شرکت‌پذیری: به ازای هر $u, v, w \in V$ و هر $a, b \in R$:

$$(u+v)+w = u+(v+w)$$

$$(ab)v = a(bv)$$

۳. عضو خنثی جمعی: عضوی مانند ۰ در V وجود داشته باشد به طوری که به ازای هر $v \in V$ داشته باشیم $v + ۰ = v$.

۴. وارون جمعی: به ازای هر $v \in V$ عضوی مانند $w \in V$ باشد به طوری که $v + w = ۰$.

۵. عضو خنثی ضربی: عضوی مانند ۱ در V وجود داشته باشد به طوری که به ازای هر $v \in V$ داشته باشیم $۱ \cdot v = v$.

۶. توزیع‌پذیری: به ازای هر $u, v \in V$ و $a, b \in R$:

$$(a+b)v = av + bv$$

$$a(u+v) = au + av$$

توجه کنید که فضای خطی در تعریف فوق، فضای خطی روی اعداد حقیقی یا همان «فضای خطی حقیقی» است. می‌توان به طور مشابه، فضای خطی را روی اعداد مختلط C نیز تعریف کرد که آن را فضای خطی **مختلط** می‌نامند. برای تشکیل چنین فضایی، کافی است که اسکالرها را از C انتخاب کنیم.

مثال: فرض کنید $A \in M_n(R)$ باشد. ستون‌های A را با $A_۱ \cdots A_n$ نمایش می‌دهیم. قرار دهید:

$$V = \{c_۱ A_۱ + \cdots + c_n A_n \mid c_i \in R, ۱ \leq i \leq n\}$$

به راحتی می‌توان دید که V یک فضای خطی است. فضای خطی V را فضای **ستونی** ماتریس A می‌نامیم و با $C(A)$ نمایش می‌دهیم.

تعریف زیرفضای خطی: فرض کنید V یک فضای خطی باشد. زیرفضای خطی، زیرمجموعه‌ای ناتهی از V مانند W است، به طوری که در شرایط فضای خطی صدق کند، یعنی:

۱. اگر هر بردار x و y از زیرفضای W را با هم جمع کنیم ، $x+y$ در W باشد.

۲. اگر v بردار دلخواهی در W و c اسکالری دلخواه باشد cv در W باشد.

یادداشت: در مثال قبل $C(A) \subseteq R^m$ و $C(A)$ فضای خطی است. بنابراین انتظار داریم $C(A)$ به عنوان زیرمجموعه‌ای از R^m، یک زیرفضای خطی از R^m باشد.

نکته: اگر V فضای خطی و $W \subseteq V$ زیرفضای خطی باشد آنگاه $\cdot \in W$.

مثال: فرض کنید $S_۱ = \{(x,y) \in R^۲ | x \geq \cdot, y \geq \cdot\}$. $S_۱$ زیرفضا نیست، زیرا:

$$v = \begin{bmatrix} ۱ \\ ۱ \end{bmatrix}, c = -۱ \quad \Rightarrow \quad c \cdot v \notin S_۱$$

حال به مجموعه‌ی $S_۱$ ربع سوم صفحه‌ی مختصات را نیز اضافه می‌کنیم، یعنی:

$$S_۲ = \{(x,y) \in R^۲ | x \geq \cdot, y \geq \cdot \quad or \quad x < \cdot, y < \cdot\}$$

در این صورت:

$$v_۱ = \begin{bmatrix} ۱ \\ ۲ \end{bmatrix}, v_۲ = \begin{bmatrix} -۲ \\ -۱ \end{bmatrix} \quad \Rightarrow \quad v_۱ + v_۲ = \begin{bmatrix} -۱ \\ ۱ \end{bmatrix} \notin S_۲$$

پس $S_۲$ نیز زیرفضا نیست.

توجه: کوچک‌ترین زیرفضای شامل $S_۱$، خود $R^۲$ است. همچنین $S_۱$ زیرمجموعه‌ای از $R^۲$ است که تحت جمع بسته بوده، ولی تحت ضرب اسکالر بسته نیست.

مثال: کوچک‌ترین زیرفضای $M_n(R)$ که شامل همه‌ی ماتریس‌های متقارن و همه‌ی ماتریس‌های مثلثی است را بیابید.

حل:

$$W = \{\alpha S + \beta L | S \text{متقارن}, L \text{پایین‌مثلثی}, \alpha, \beta \in R\}$$

فرض کنید $A \in M_n(R)$. در این صورت قرار دهید:

$$S_{ij} = \begin{cases} A_{ij} & i < j \\ \cdot & i = j \\ A_{ji} & i > j \end{cases} \quad \Rightarrow \quad (A-S)_{ij} = \begin{cases} \cdot & i \leq j \\ a_{ij} - a_{ji} & i > j \end{cases}$$

بنابراین $A = (A-S) + S$ که در آن $A - S$ پایین‌مثلثی و S متقارن است. پس کوچک‌ترین زیرفضای شامل همه‌ی ماتریس‌های متقارن و پایین‌مثلثی، همان $M_n(R)$ است.

مثال: فرض کنید $A \in M_n(R)$ باشد؛ آن‌گاه، زیرمجموعه‌ی $W = \{x \in R^n | Ax = \cdot\} \subseteq R^n$ یک زیرفضای خطی است که آن را **فضای پوچ** A می‌نامیم و با نماد $N(A)$ نمایش می‌دهیم.

گزاره: فرض کنید $A \in M_n(R)$ باشد؛ در این صورت،

۱. A وارون‌پذیر است اگر و تنها اگر $N(A) = \{\cdot\}$

۲. A وارون‌پذیر است اگر و تنها اگر $C(A) = R^n$

برهان:

۱. برای اثبات جهت اول، فرض می‌کنیم $Ax = \cdot$؛ نتیجه می‌گیریم $x = A^{-۱}\cdot = \cdot$ و در نتیجه $N(A) = \{\cdot\}$. برای اثبات جهت دوم، تجزیه‌ی LU برای A را در نظر می‌گیریم. فرض می‌کنیم $LUx = \cdot$؛ ماتریس L نماینده‌ی اعمال سطری و در نتیجه وارون‌پذیر است، پس می‌توانیم نتیجه بگیریم که $Ux = \cdot$. ادعا می‌کنیم که عناصر روی قطر اصلی U ناصفر هستند.

$$Ux = \begin{bmatrix} u_{۱۱} & & * \\ & \ddots & \\ \cdot & & u_{nn} \end{bmatrix} x = \cdot$$

به ازای هر i از مجموعه‌ی $\{۱, \cdots, n\}$، داریم $u_{ii} = \cdot$؛ چرا که در غیر این صورت، با تبدیل U به ماتریس تحویل‌یافته‌ی سطری پلکانی U' خواهیم دید که دستگاه $U'x = \cdot$ و در نتیجه $Ux = \cdot$ متغیر آزاد دارد و در نتیجه،

$$N(A) = N(U) \neq \cdot$$

حال ثابت می‌کنیم هر ماتریس بالامثلثی با درایه‌های قطر اصلی ناصفر، وارون‌پذیر است. با

استقرا روی n ثابت می‌کنیم. به ازای شرط پایه‌ی $n = ۲$،

$$U = \begin{bmatrix} u_۱ & a \\ ۰ & u_۲ \end{bmatrix}$$

و وارون آن،

$$\begin{bmatrix} \frac{۱}{u_۱} & \frac{-a}{u_۱ u_۲} \\ ۰ & \frac{۱}{u_۲} \end{bmatrix}$$

است.

فرض استقرا: فرض کنید هر ماتریس بالامثلثی $U_{n-۱}$ در $M_{n-۱}(R)$ با درایه‌های قطر اصلی ناصفر، وارون‌پذیر است.

حکم استقرا: برای U_n در $M_n(R)$ با شرایط مفروض، وارون‌پذیری را ثابت می‌کنیم. قرار دهید:

$$U_n = \begin{bmatrix} a & C^T \\ ۰ & U_{n-۱} \end{bmatrix}$$

وارون U_n، ماتریس

$$U_n{}^{-۱} = \begin{bmatrix} \frac{۱}{a} & x^T \\ ۰ & U_{n-۱}{}^{-۱} \end{bmatrix}$$

است که در آن،

$$x^T = -\frac{۱}{a} C^T U_{n-۱}{}^{-۱}$$

درنتیجه اگر $N(A) = \{۰\}$ آن‌گاه در $A = LU$، ماتریس U وارون‌پذیر است، پس داریم

$$A^{-۱} = U^{-۱} L^{-۱}$$

۲. برای اثبات جهت اول، فرض می‌کنیم A وارون‌پذیر باشد. ستون‌های A را با A_i ها نمایش می‌دهیم. آن‌گاه،

$$C(A) = \{c_۱ A_۱ + \cdots + c_n A_n | c_i \in R, ۱ \leq i \leq n\}$$

زیرمجموعه‌ای از R^n است. فرض کنید b عضوی از R^n باشد. باید c_i های حقیقی بیابیم که $c_1A_1 + \cdots + c_nA_n = b$. از طرفی،

$$c_1A_1 + \cdots + c_nA_n = \begin{bmatrix} A_1 & \cdots & A_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = b$$

حال، چون A وارون‌پذیر است، $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = A^{-1}b$ و در نتیجه، $C(A) = R^n$.

برای اثبات جهت دوم، می‌دانیم $C(A) = R^n$، در نتیجه به ازای هر b که در R^n است، بردار $c \in R^n$ وجود دارد که $Ac = b$ باشد؛ بنابراین به ازای هر i که در نظر بگیریم، $Ax = e_i$ جواب دارد. فرض کنید c_i جواب $Ax = e_i$ باشد و قرار دهید $c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$. در این صورت، $AC = I$ و اصطلاحا می‌گوییم A وارون راست دارد. حال ثابت می‌کنیم که C و در نتیجه A وارون‌پذیرند. برای اثبات، از مورد ۱ استفاده می‌کنیم و نشان می‌دهیم $N(C) = \{0\}$. فرض کنید $Cx = 0$؛ در نتیجه $ACx = Ix = x = 0$ و پس x صفر است، یعنی $N(C) = \{0\}$ و در نتیجه C وارون‌پذیر است، و $AC = I$ پس $CA = I$ و A نیز وارون‌پذیر است.

مثال: فرض کنید W_1 و W_2 دو زیرفضا از فضای برداری V باشند. آیا $W_1 \cup W_2$ زیرفضاست؟

پاسخ: فرض کنید $W_1 = \{(x, 0) | x \in R\}$ و $W_2 = \{(0, y) | y \in R\}$. در این صورت $W_1 \cup W_2$ زیرفضا نیست، زیرا:

$$(1, 0) \in W_1 \cup W_2 \quad , \quad (0, 1) \in W_1 \cup W_2$$

ولی:

$$(1, 0) + (0, 1) = (1, 1) \notin W_1 \cup W_2$$

گزاره: اگر W_1 و W_2 دو زیرفضا از V باشند که $W_1 \cup W_2$ زیرفضا شود، آن‌گاه $W_1 \subseteq W_2$ یا $W_2 \subseteq W_1$.

برهان: برای اثبات این موضوع، از برهان خلف کمک می‌گیریم. فرض کنید $v_1 \in W_1 \setminus W_2$

و $W_۱ \setminus W_۲ \ni v_۲$؛ در نتیجه، $v_۱, v_۲ \in W_۱ \cup W_۲$. چون $W_۱ \cup W_۲$ زیرفضا است، پس $v_۱ + v_۲ \in W_۱ \cup W_۲$ و در نتیجه، $v_۱ + v_۲ \in W_۱$ یا $v_۱ + v_۲ \in W_۲$. اگر $v_۱ + v_۲ \in W_۲$، آنگاه $v_۱ = v_۱ + v_۲ - v_۲ \in W_۲$ که تناقض است. به طریق مشابه از $v_۱ + v_۲ \in W_۲$ نیز به تناقض می‌رسیم؛ پس فرض خلف باطل است و $W_۱ \subseteq W_۲$ و یا $W_۲ \subseteq W_۱$.

با اجتماع تعدادی زیرفضا، نمی‌توان همواره زیرفضا ساخت، اما با «جمع» آن‌ها به تعبیر زیر می‌توان زیرفضا ساخت:

تعریف: فرض کنید $W_۱, \cdots, W_k$ زیرفضاهایی از فضای برداری V باشند. **مجموع** این زیرفضاها را به شکل زیر تعریف می‌کنیم:

$$W_۱ + \cdots + W_k := \{v_۱ + \cdots + v_k | \forall i, v_i \in W_۱\}$$

گزاره: اگر $W_۱, \cdots, W_k$ زیرفضاهایی از فضای برداری V باشند، $W_۱ + \cdots + W_k$ نیز یک زیرفضای برداری است.

برهان: ابتدا باید ثابت کنیم که عدد صفر در این زیرفضا وجود دارد.

$$۰ = ۰ + \cdots + ۰ \in W_۱ + \cdots + W_k$$

حال فرض کنید $w_۱ + \cdots + w_k, v_۱ + \cdots + v_k \in W_۱ + \cdots + W_k$ و $c \in R$؛ در این صورت:

$$c(v_۱ + \cdots + v_k) + (w_۱ + \cdots + w_k) = (cv_۱ + w_۱) + \cdots + (cv_k + w_k) \in W_۱ + \cdots + W_k \ .$$

حل $Ax = b$ و $Ax = ۰$:
به ازای ماتریس $A \in M_{mn}(R)$، حل دستگاه $Ax = ۰$ معادل یافتن زیرفضای $N(A)$ یا فضای پوچ A است.
دستگاه $Ax = b$ جواب دارد، اگر و تنها اگر $b \in C(A)$.
اگر A مربعی باشد ($m = n$)، در صورت امکان با اعمال الگوریتم حذف گاوسی، A تجزیه‌ی LU دارد؛ بنابراین

$$LUx = Ax = b \quad \Rightarrow \quad Ux = L^{-۱}b \ .$$

در غیر این صورت، با جابه‌جایی سطرها امکان اعمال الگوریتم مهیا می‌شود و PA تجزیه‌ی

LU خواهد داشت؛ در نتیجه:

$$Ax = b \quad \Rightarrow \quad PAx = Pb \quad \Rightarrow \quad LUx = Pb \quad \Rightarrow \quad Ux = L^{-۱}Pb\,.$$

چون U ماتریسی بالامثلثی است، امکان یافتن x وجود دارد. حال باید فرایندی مشابه را برای حل دستگاه $Ax = b$ ــ وقتی A لزوماً مربعی نیست ــ پیاده کنیم. مشابه حالت قبل، از اعمال عملیات سطری مقدماتی، ماتریس U را به دست می‌آوریم.

مثال: فرض کنید $A \in M_{۳\times۴}(R)$ به طوری که:

$$A = \begin{bmatrix} ۱ & ۳ & ۳ & ۲ \\ ۲ & ۶ & ۹ & ۷ \\ -۱ & -۳ & ۳ & ۴ \end{bmatrix}\,.$$

درایه‌ی $a_{۱۱} = ۱ \neq ۰$، در نتیجه به عنوان درایه‌ی محوری منظور می‌گردد. با استفاده از عملیات سطری مقدماتی، درایه‌های $a_{۲۱}$ و $a_{۳۱}$ را صفر می‌کنیم. توجه کنید که این اعمال سطری مقدماتی معادل ماتریس‌هایی مقدماتی و وارون‌پذیر هستند؛ به عبارت دیگر، برای صفر کردن $a_{۲۱}$ و $a_{۳۱}$ دو ماتریس مقدماتی وارون‌پذیر سه در سه $E_۱$ و $E_۲$ در A ضرب می‌شود، به طوری که:

$$E_۱E_۲A = \begin{bmatrix} ۱ & ۳ & ۳ & ۲ \\ ۰ & ۰ & ۳ & ۳ \\ ۰ & ۰ & ۶ & ۶ \end{bmatrix}\,.$$

در مرحله‌ی بعد، از درایه‌ی سطر دوم و ستون سوم به عنوان درایه‌ی محوری استفاده می‌کنیم و فرایند قبل را ادامه می‌دهیم تا به حاصل زیر برسیم:

$$E'A = \begin{bmatrix} ۱ & ۳ & ۳ & ۲ \\ ۰ & ۰ & ۳ & ۳ \\ ۰ & ۰ & ۰ & ۰ \end{bmatrix}$$

این ماتریس را با U نمایش می‌دهیم و ماتریس «پلکانی» یا «سطری پلکانی» می‌نامیم. توجه کنید که $E'A = U$، که در آن، E' ماتریسی وارون‌پذیر است.

برای سادگی و محاسبه‌ی سریع‌تر دستگاه، درایه‌های محوری را تبدیل به یک می‌کنیم و درایه‌های

بالای درایه‌های محوری را نیز صفر می‌کنیم. با انجام فرایند فوق، داریم:

$$E''U = (E''E')A = EA = \begin{bmatrix} ۱ & ۳ & ۰ & -۱ \\ ۰ & ۰ & ۱ & ۱ \\ ۰ & ۰ & ۰ & ۰ \end{bmatrix} = R$$

ماتریس حاصل را ماتریس «تحویل‌یافته‌ی سطری پلکانی» می‌نامیم و با R نمایش می‌دهیم. توجه کنید که R حاصل‌ضرب ماتریسی وارون‌پذیر در A است.

بنابراین، حل $Ax = b$ معادل است با حل $Rx = Eb$.

تمرین: فرض کنید

$$A = \begin{bmatrix} ۱ & ۳ & ۳ & ۲ \\ ۲ & ۶ & ۹ & ۷ \\ -۱ & -۳ & ۳ & ۴ \end{bmatrix}$$

الف) $N(A)$ را به دست آورید.
ب) دستگاه $Ax = b$ به ازای چه $b \in R^۳$ جواب دارد؟
ج) جواب دستگاه $Ax = b$ را در صورت وجود محاسبه کنید.

پاسخ:

الف)

$$N(A) = \left\{ v \begin{bmatrix} -۳ \\ ۱ \\ ۰ \\ ۰ \end{bmatrix} + y \begin{bmatrix} ۱ \\ ۰ \\ -۱ \\ ۱ \end{bmatrix} \middle| v, y \in \mathbb{R} \right\}$$

ب)

$$C(A) = \left\{ v \begin{bmatrix} ۱ \\ ۲ \\ -۱ \end{bmatrix} + y \begin{bmatrix} ۲ \\ ۷ \\ ۴ \end{bmatrix} \middle| v, y \in \mathbb{R} \right\}$$

ج) جواب دستگاه، شامل تمام x هایی است که در معادله‌ی ماتریسی زیر، صدق کنند:

$$\begin{bmatrix} ۱ & ۳ & ۰ & ۱ \\ ۰ & ۰ & ۱ & ۱ \\ ۰ & ۰ & ۰ & ۰ \end{bmatrix} x = E'Eb$$

توجه کنید که ضرب ماتریس‌های مقدماتی، باید در هر دو طرف تساوی انجام شود. **تعریف:** فرض کنید V یک زیرفضای برداری باشد و $\emptyset \neq S \subseteq V$. در این صورت زیرفضای تولید شده توسط زیرمجموعه‌ی ناتهی S را برابر مجموعه‌ی تمام ترکیبات خطی عناصر S تعریف می‌کنیم و با $span(S)$ نمایش می‌دهیم. به عبارت دیگر:

$$span(S) = \{c_۱v_۱ + \cdots + c_nvn | v_i \in S, c_i \in R\} .$$

مثال: فرض کنید:

$$A = \begin{bmatrix} ۱ & ۳ & ۳ & ۲ \\ ۲ & ۶ & ۹ & ۷ \\ -۱ & -۳ & ۳ & ۴ \end{bmatrix} .$$

فضای ستونی A، فضای تولیدشده توسط ستون‌های A است. اگر ستون iام A را با A_i نمایش دهیم:

$$\begin{aligned} C(A) &= span(\{A_۱, A_۲, A_۳, A_۴\}) \\ &= \left\{ c_۱ \begin{bmatrix} ۱ \\ ۲ \\ -۱ \end{bmatrix} + c_۲ \begin{bmatrix} ۳ \\ ۶ \\ -۳ \end{bmatrix} + c_۳ \begin{bmatrix} ۳ \\ ۹ \\ ۳ \end{bmatrix} + c_۴ \begin{bmatrix} ۲ \\ ۷ \\ ۴ \end{bmatrix} \middle| c_i \in R \right\} \\ &= \left\{ c_۱ \begin{bmatrix} ۱ \\ ۲ \\ -۱ \end{bmatrix} + c_۲ \begin{bmatrix} ۲ \\ ۷ \\ ۴ \end{bmatrix} \middle| c_۱, c_۲ \in R \right\} \\ &= span(\{A_۱, A_۴\}) . \end{aligned}$$

تعریف: فرض کنید V یک فضای برداری حقیقی باشد. مجموعه‌ی ناتهی S از V را وابسته‌ی خطی گوییم، هرگاه وجود داشته باشد $v_1, \cdots, v_n \in S$ و $c_1, \cdots, c_n \in R$ که $c_1v_1 + \cdots + c_nv_n = 0$ و حداقل یکی از c_i ها ناصفر باشد. S را مستقل خطی می‌گوییم، هرگاه وابسته‌ی خطی نباشد.

نکته: فرض کنید $S = \{v_1, \cdots, v_n\}$. برای تشخیص وابسته بودن اعضای S یا مستقل بودن آن‌ها باید معادله‌ی

$$c_1v_1 + \cdots + c_nv_n = 0$$

را تشکیل دهیم که معادل است با $AC = 0$ که در آن $A = \begin{bmatrix} v_1 \cdots v_n \end{bmatrix}$ و $C = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$.

بنابراین بردارهای $v_1 \cdots v_n$ مستقل خطی‌اند، اگر و تنها اگر $N(A) = \{0\}$.

مثال: ثابت کنید بردارهای زیر وابسته‌ی خطی هستند.

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad v_2 = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix} \quad v_3 = \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix}$$

پاسخ: کافی است بردارها را به صورت ستونی در یک ماتریس بنویسیم و سپس با نوشتن $AC = 0$ نتیجه بگیریم که برای cها، می‌توانیم مقداری غیر از صفر داشته باشیم.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 7 & 9 \\ -1 & 4 & 3 \end{bmatrix}$$

پس از تبدیل این ماتریس به ماتریس بالامثلثی داریم:

$$EA = U = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

حال دستگاه زیر را در نظر بگیرید:

$$\begin{bmatrix} ۱ & ۲ & ۳ \\ ۰ & ۱ & ۱ \\ ۰ & ۰ & ۰ \end{bmatrix} \begin{bmatrix} c_۱ \\ c_۲ \\ c_۳ \end{bmatrix} = \begin{bmatrix} ۰ \\ ۰ \\ ۰ \end{bmatrix} .$$

از سطر آخر داریم:

$$c_۱ \times ۰ + c_۲ \times ۰ + c_۳ \times ۰ = ۰ .$$

بنابراین متوجه می‌شویم که برای $c_۳$ بی‌نهایت مقدار داریم و لزومی ندارد که $c_۳ = ۰$ باشد.

قضیه: اگر $V = span(\{v_۱, \cdots, v_n\})$، هر مجموعه‌ی مستقل خطی از V، حداکثر n عضو دارد.

برهان: به برهان خلف، فرض کنید $\{w_۱, \cdots, w_k\}$ یک مجموعه‌ی مستقل خطی از V باشد. هم‌چنین، فرض کنید $k > n$. چون $V = span(\{v_۱, \cdots, v_n\})$، بنابراین به ازای هر $۱ \leq j \leq k$:

$$w_j = \sum_{i=۱}^{n} c_{ij} v_i$$

قرار دهید $A = [c_{ij}]$ و دستگاه $Ax = ۰$ را در نظر بگیرید.

$$Ax = \begin{bmatrix} c_{۱۱} & \cdots & c_{nk} \\ \vdots & & \vdots \\ c_{n۱} & \cdots & c_{nk} \end{bmatrix}_{nk} \begin{bmatrix} x_۱ \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} ۰ \\ \vdots \\ ۰ \end{bmatrix}$$

که معادل است با ترکیب خطی $x_۱ w_۱ + \cdots + x_k w_k = ۰$. چون $\{w_۱, \cdots, w_k\}$ مجموعه‌ای مستقل است، $Ax = ۰$ فقط یک جواب دارد و آن $x = ۰$ است؛ معادلاً، $N(A) = \{۰\}$. از طرفی چون $k > n$، پس متغیر آزاد وجود دارد؛ لذا دستگاه $Ax = ۰$ جواب ناصفر نیز دارد و در نتیجه $N(A) \neq \{۰\}$. به تناقض رسیدیم، پس فرض خلف $k > n$ باطل بوده و حکم ثابت می‌شود.

تعریف: اگر V فضای برداری حقیقی باشد، یک زیرمجموعه‌ی مستقل خطی از V که V را تولید می‌کند، پایه برای V نامیده می‌شود.

مثال:

$$A=\begin{bmatrix} ۱ & ۲ & -۱ & ۰ & ۱ & ۰ \\ -۱ & -۱ & ۲ & -۳ & ۱ & ۰ \\ ۱ & ۱ & -۲ & ۰ & ۰ & ۲ \\ ۰ & ۰ & ۰ & ۳ & ۱ & -۲ \end{bmatrix}$$

فضای تصویر A را با V نشان می‌دهیم. در این صورت:

$$\begin{aligned} V &= \left\{ x_۱\underbrace{\begin{bmatrix} ۱ \\ -۱ \\ ۱ \\ ۰ \end{bmatrix}}_{v_۱} + x_۲\underbrace{\begin{bmatrix} ۲ \\ -۱ \\ ۱ \\ ۰ \end{bmatrix}}_{v_۲} + x_۳\underbrace{\begin{bmatrix} -۱ \\ ۲ \\ -۲ \\ ۰ \end{bmatrix}}_{v_۳} + x_۴\underbrace{\begin{bmatrix} ۰ \\ -۳ \\ ۰ \\ ۳ \end{bmatrix}}_{v_۴} + x_۵\underbrace{\begin{bmatrix} ۱ \\ ۱ \\ ۰ \\ ۱ \end{bmatrix}}_{v_۵} + x_۶\underbrace{\begin{bmatrix} ۰ \\ ۰ \\ ۲ \\ -۲ \end{bmatrix}}_{v_۶} \,\middle|\, x_i \in \mathbb{R} \right\} \\ &= C(A) \\ &= span(\{v_۱, v_۲, v_۳, v_۴, v_۵, v_۶\}) \\ &= span(\{v_۱, v_۲, v_۴, v_۵\}). \end{aligned}$$

تعریف: فضای برداری V را دارای بعد متناهی گوییم، هرگاه دارای پایه‌ای متناهی باشد.

قضیه: اگر V فضای برداری با بعد متناهی باشد، در این صورت هر دو پایه‌ی V به تعداد مساوی عضو دارند.

برهان: فرض کنید $B=\{v_۱,\cdots,v_n\}$ و $B'=\{w_۱,\cdots,w_m\}$ دو پایه برای V باشند. چون $V=span(\{v_۱,\cdots,v_n\})$ و B' یک مجموعه‌ی مستقل است، بنا به قضیه‌ی قبل، $m\leq n$. به طریق مشابه، چون $V=span(\{w_۱,\cdots,w_m\})$ و B یک مجموعه‌ی مستقل است، $n\leq m$ و در نتیجه $n=m$.

تعریف: بعد یک فضای برداری با بعد متناهی، برابر تعداد اعضای پایه‌ی آن تعریف می‌شود. بعد V را با $dim(V)$ نشان می‌دهیم.

نکته: فرض کنید $A\in M_{mn}(R)$ و می‌خواهیم $N(A)$ را محاسبه کنیم؛ برای این کار، ماتریس تحویل یافته‌ی سطری پلکانی A را محاسبه می‌کنیم. فرض کنید این ماتریس، r سطر ناصفر(معادلاً r درایه‌ی محوری) داشته باشد. اگر ستون‌های محوری R را با اندیس‌های $j_۱,\cdots,j_r$ نمایش دهیم، متغیرهای مربوط به سایر ستون‌ها، متغیرهای آزاد بوده و متغیرهای محوری به صورت زیر هستند:

$$
\begin{cases}
x_{j_r} = -\sum\limits_{i=j_r+1}^{n} c_{ri}x_i \\
x_{j_{r-1}} = -\sum\limits_{i=j_{ri}^{+1}}^{n} c_{(r-1)i}x_i \\
\vdots \\
x_{j_1} = -\sum\limits_{i=j_1+1}^{n} c_{1i}x_i
\end{cases}
\tag{۱.۳}
$$

به وضوح $N(A) = span\Big(\Big\{e_j | j \in \{1, \cdots n\} \setminus \{j_1, \cdots, j_r\}\Big\}\Big)$؛ در نتیجه،

$$
dim(N(A)) = n - r \ .
$$

می‌دانیم اگر $V = span(\{v_1, \cdots, v_n\})$، به ازای هر بردار $v \in V$ وجود دارد $c_1, \cdots, c_n \in R$ به طوری که $v = c_1v_1 + \cdots + c_nv_n$. سؤالی که ممکن است پیش بیاید، این است که آیا ترکیب خطی ارائه شده برای v، یکتاست؟

فرض کنید $v = d_1v_1 + \cdots + d_nv_n$. در نتیجه، داریم:

$$
c_1v_1 + \cdots + c_nv_n = d_1v_1 + \cdots + d_nv_n
$$

$$
\Rightarrow (c_1 - d_1)v_1 + \cdots + (c_n - d_n)v_n = 0 \quad .
$$

برای این‌که این تجزیه (یا ترکیبات خطی برای v بر اساس بردارهای تولیدکننده) یکتا باشد، بایستی به ازای هر $1 \leq i \leq n$ داشته باشیم $c_i = d_i$؛ پس به استقلال خطی بردارهای $v_1, \cdots, v_n$ نیاز داریم. بنابراین، اگر برای v یک پایه مانند $B = \{v_1, \cdots v_k\}$ داشته باشیم، چنین فرض می‌کنیم که اعضای پایه مستقل خطی هستند و در نتیجه نمایش ترکیب خطی هر بردار v نسبت به آن پایه، یکتاست. بردار $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ (اگر $c_1v_1 + \cdots + c_nv_n$) را مختصات بردار v در پایه‌ی $B = \{v_1 \cdots v_k\}$ می‌نامند.

توجه کنید که برای آن که مختصات یک بردار در پایه، خوش تعریف باشد، باید ترتیبی روی عناصر

پایه لحاظ کنیم.

مختصات یک بردار در یک پایه‌ی مرتب

تعریف: فرض کنید V یک فضای خطی و $B = \{v_۱ \cdots v_n\}$ یک پایه برای V باشد. در این صورت، B را یک پایه‌ی **مرتب** برای V گویند، هرگاه ترتیب v_i ها در B در نظر گرفته شود.

نکته: اگر $v \in V$، وجود دارند $c_۱, \cdots, c_n \in R$ هایی، به طوری که

$$v = c_۱v_۱ + \cdots + c_nv_n$$

و c_i ها یکتا هستند، چون v_i ها مستقل خطی هستند.

تعریف: بردار $c = \begin{bmatrix} c_۱ \\ \vdots \\ c_n \end{bmatrix}$ را بردار مختصات $v = c_۱v_۱ + \cdots + c_nv_n$ در پایه‌ی B گوییم و با نماد $[v]_B$ نشان می‌دهیم.

سؤال: اگر پایه را عوض کنیم و پایه‌ی جدید B' باشد، $[v]_B$ و $[v]_{B'}$ چه ارتباطی با یک‌دیگر دارند؟

اگر $B' = \{v'_۱, \cdots, v'_n\}$ پایه‌ی جدیدی برای V باشد و $v = c'_۱v'_۱ + \cdots + c'_nv'_n$ ، آن‌گاه:

$$[v]_{B'} = \begin{bmatrix} c'_۱ \\ \vdots \\ c'_n \end{bmatrix}$$

از طرفی $v'_i \in V$؛ در نتیجه وجود دارد $p_{ij} \in R$ به طوری که برای هر $i \in \{۱, \cdots, n\}$:

$$v'_j = \sum_{i=۱}^{n} p_{ij}v_i$$

داریم:

$$v = \sum_{j=۱}^{n} c'_jv'_j$$

$$\Rightarrow v = \sum_{j=۱}^{n} c'_j \sum_{i=۱}^{n} p_{ij}v_i$$

$$\Rightarrow v = (\sum_{j=1}^{n} c'_j p_{1j})v_1 + \cdots + (\sum_{j=1}^{n} c'_j p_{nj})v_n$$

$$\Rightarrow [v]_B = \begin{bmatrix} \sum_{j=1}^{n} c'_j p_{1j} \\ \vdots \\ \sum_{j=1}^{n} c'_j p_{nj} \end{bmatrix} = \underbrace{\begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}}_{P} \begin{bmatrix} c'_1 \\ \vdots \\ c'_n \end{bmatrix} = P[v]_{B'}$$

$$[v]_B = P[v]_{B'}$$

در واقع اگر p_j ستون jام P باشد، آنگاه $p_j = [v'_j]_B$.

نکته: توجه کنید که P وارون‌پذیر است؛ چرا که اگر $v \in N(P)$، آنگاه $P[v]_{B'} = 0$ و بنابراین $[v]_B = 0$، در نتیجه $v = 0$.

گزاره: فرض کنید V فضای خطی با بعد n باشد و B و B' دو پایه برای V باشند. در این صورت ماتریس یکتای P وجود دارد که به ازای هر بردار $v \in V$، داشته باشیم

$$[v]_B = P[v]_B \quad .$$

برهان: تنها لازم است که یکتایی P را ثابت کنیم. فرض کنید

$$[v]_B = P[v]_{B'} \quad , \quad [v]_B = Q[v]_{B'}$$

به ازای هر $v \in V$و P, Q وارون‌پذیر باشند. آنگاه:

$$(P - Q)[v]_{B'} = 0$$

چون $[v]_{B'}$ همه‌ی بردارهای R^n را می‌سازد، به ازای هر j، $1 \leq j \leq n$، $(P - Q)e_j = 0$. از طرفی $(P - Q)e_j$ ستون j ام ماتریس $(P - Q)$ است؛ پس $P - Q = 0$ و در نتیجه $P = Q$.

مثالی از فضای خطی با بعد نامتناهی

فرض کنید $R[x]$ فضای خطی همه‌ی چندجمله‌ای‌ها با ضرایب حقیقی باشد. نشان می‌دهیم که $R[x]$ یک فضای خطی با بعد نامتناهی است.

برای نشان دادن این حکم، از برهان خلف استفاده می‌کنیم. فرض کنید بعد $R[x]$ متناهی باشد. بنابراین $R[x]$ دارای یک پایه‌ی متناهی است. فرض کنید $\{f_1, \cdots, f_n\}$ پایه‌ی $R[x]$ باشد. درجه‌ی

هر چندجمله‌ای f_i، $1 \leq i \leq n$، را با $deg\, f_i$ نمایش می‌دهیم. قرار دهید $m = max\, deg\, f_i$. در این صورت، چندجمله‌ای x^{m+1} را نمی‌توان بر حسب ترکیب خطی از چندجمله‌ای‌های $f_1, \cdots, f_n$ که فرض شده بود پایه‌ی $R[x]$ اند، نوشت که این موضوع، متناقض با فرض می‌باشد. در نتیجه، $R[x]$ پایه‌ی متناهی ندارد.

توجه: فرض کنید V یک فضای خطی و $S \subseteq V$ یک زیرمجموعه‌ی مستقل خطی باشد. اگر $v \in V - span(S)$ باشد (یعنی v برداری در V باشد که در فضای تولید شده توسط S قرار نداشته باشد)، آنگاه $S \cup \{v\}$ مجموعه‌ی مستقل خطی است. (چرا؟)

زیرا: به برهان خلف، اگر $S \cup \{v\}$ مستقل خطی نباشد، آنگاه v را می‌توان به صورت ترکیب خطی از اعضای S نوشت(زیرا S مستقل خطی است) که تناقض است، زیرا فرض شده بود که $v \notin span(S)$.

توجه: اگر $dim\ V < \infty$ و $W \subseteq V$ یک زیرفضای خطی از فضای خطی V باشد، آنگاه هر زیرمجموعه‌ی مستقل خطی از W متناهی بوده و قسمتی از یک پایه برای فضای خطی V است. (چرا؟)

زیرا: اگر $dim\ V = n$، هر زیرمجموعه‌ی مستقل خطی از V حداکثر n عضو دارد (با توجه به یکتایی تعداد اعضای پایه)؛ حال فرض کنید $\{v_1, \cdots, v_m\}$ یک مجموعه‌ی مستقل خطی از W باشد:

حالت اول: $V = span(\{v_1, \cdots, v_m\})$ پس $\{v_1, \cdots, v_m\}$ یک پایه برای V است.

حالت دوم: $V \neq (\{v_1, \cdots, v_m\})$؛ بنابراین $w_1 \in V - span(\{v_1, \cdots, v_m\})$. با توجه به نکته‌ی فوق مجموعه‌ی $v_1, \cdots, v_m, w_1$ مستقل خطی است. حال، حالت اول را چک می‌کنیم. اگر

$$v = span(\{v_1, \cdots, v_m, w_1\})$$

بنابراین $v = span(\{v_1, \cdots, v_m, w_1\})$ پایه‌ای برای V است و الگوریتم پیدا کردن پایه برای V به پایان می‌رسد، وگرنه حالت دوم $V \neq span(\{v_1, \cdots, v_m, w_1\})$ رخ می‌دهد. پس $w_2 \in V - span(\{v_1, \cdots, v_m, w_1\})$ در نظر می‌گیریم و مراحل را ادامه می‌دهیم. چون بعد V متناهی است، این الگوریتم حداکثر $n - m$ مرحله به پایان می‌رسد.

توجه: فرض کنید V فضای خطی با بعد متناهی است ($dim\, v = n$). با الگوریتم فوق، می‌توانیم یک پایه برای V بیابیم. برای انجام این کار، $0 \neq v_1 \in V$ را در نظر بگیرید و قرار دهید $W = span(\{v_1\})$. با اضافه کردن بردار به $\{v_1\}$ با استفاده از الگوریتم فوق، می‌توانیم پایه‌ای برای V بیابیم.

توجه: فرض کنید W_1 و W_2 دو زیرفضای خطی از فضای V با بعد متناهی باشند. آن‌گاه، $W_1 + W_2$ نیز زیرفضایی خطی با بعد متناهی است و

$$dim(W_1 + W_2) = dimW_1 + dimW_2 - dim(W_1 \cap W_2)$$

زیرا: ابتدا نشان می‌دهیم که $W_1 \cap W_2$ خود یک زیرفضاست. چون W_1 و W_2 زیرفضا هستند، بردار صفر در آن‌ها قرار دارد و در نتیجه بردار صفر در اشتراک آن‌ها نیز واقع است. حال اگر دو بردار a و b در اشتراک دو زیرفضا باشند، پس در هر دو قرار دارند، در نتیجه جمع آن‌ها و نیز ضرب اسکالر عدد در آن‌ها در هر دو زیرفضا (و معادلا، در اشتراک آن‌ها) قرار دارد.

حال ثابت می‌کنیم که $W_1 + W_2$ نیز خود یک زیرفضاست. مشخصا $0 = 0 + 0$، در نتیجه $W_1 + W_2$ شامل بردار صفر است. اگر دو بردار هم‌چون بردارهای a و b در $W_1 + W_2$ باشند، می‌توان نوشت:

$$a = x + y,\ b = z + t,\ x \in W_1,\ y \in W_2,\ z \in W_1,\ t \in W_2$$

و آن‌گاه، داریم:

$$a + b = (x + y) + (z + t) = (x + z) + (y + t)$$

و چون W_1 و W_2 زیرفضا هستند،

$$(x + z) \in W_1,\ (y + t) \in W_2 \rightarrow (a + b) \ \in \ (W_1 + W_2)$$

در نهایت، اگر $v \in W_1 + W_2$ و r اسکالر باشد،

$$\exists x \in W_1,\ y \in W_2 \ : \ v = x + y \rightarrow rv = r(x + y) = (rx + ry) \in (W_1 + W_2)$$

و به این شکل، زیرفضا بودن $W_1 + W_2$ ثابت می‌شود.

سپس، به اثبات اصلی می‌پردازیم؛ چون $W_1 \cap W_2 \subseteq W_1$، پس بعد $W_1 \cap W_2$ متناهی است. یک پایه برای $W_1 \cap W_2$ در نظر می‌گیریم و با استفاده از الگوریتم فوق، پایه‌هایی برای W_1 و W_2 می‌سازیم.

فرض کنید $dim(W_1 \cap W_2) = r$ و $\{p_1, \cdots, p_r\}$ یک پایه برای $W_1 \cap W_2$ باشد.

۱. با استفاده از الگوریتم فوق، مجموعه‌ی $\{p_1, \cdots, p_r\}$ را به یک پایه برای W_1 گسترش می‌دهیم:

$$\{p_1, \cdots, p_r, v_1, \cdots, v_n\} \to dimW_1 = r + n$$

۲. با استفاده از الگوریتم فوق، مجموعه‌ی $\{p_1, \cdots, p_r\}$ را به یک پایه برای W_2 گسترش می‌دهیم:

$$\{p_1, \cdots, p_r, w_1, \cdots, w_m\} \to dimW_2 = r + m$$

۳. مجموعه‌ی $\{p_1, \cdots, p_r, v_1, \cdots, v_n, w_1, \cdots, w_m\}$ پایه‌ای برای $W_1 + W_2$ است. (چرا؟)

واضح است که

$$span(\{p_1, \cdots, p_r, v_1, \cdots, v_n, w_1, \cdots, w_m\}) \subseteq W_1 + W_2$$

حال فرض کنید که $v \in W_1 + W_2$، در نتیجه $v = a + b$ به طوری که $a \in W_1$ و $b \in W_2$ و هم‌چنین

$$a = \sum_{i=1}^{r} c_i p_i + \sum_{i=1}^{n} d_i v_i$$

$$b = \sum_{i=1}^{r} c'_i p_i + \sum_{i=1}^{n} d'_i w_i$$

$$\to v = a + b = \sum_{i=1}^{r} (c_i + c'_i) p_i + \sum_{i=1}^{n} d_i v_i + \sum_{i=1}^{n} d'_i w_i$$

بنابراین

$$span(\{p_1, \cdots, p_r, v_1, \cdots, v_n, w_1, \cdots, w_m\}) = W_1 + W_2$$

حال باید نشان دهیم که این مجموعه ($\{p_1, \cdots, p_r, v_1, \cdots, v_n, w_1, \cdots, w_m\}$) مستقل خطی نیز است. برای اثبات، باید ترکیبی خطی از بردارها را به صورت زیر در نظر بگیریم:

$$\sum_{i=1}^{r} c_i p_i + \sum_{i=1}^{n} d_i v_i + \sum_{i=1}^{m} d'_i w_i = 0 \; *$$

$$\rightarrow (\sum_{i=1}^{r} c_i p_i + \sum_{i=1}^{n} d_i v_i) \in W_2 = (-\sum_{i=1}^{m} d'_i w_i) \in W_2$$

$$\rightarrow -\sum_{i=1}^{m} d'_i w_i \in W_1 \cap W_2 \rightarrow -\sum_{i=1}^{m} d'_i w_i = \sum_{i=1}^{r} t_i p_i$$

با جای‌گزینی در * داریم:

$$\sum_{i=1}^{r} c_i p_i + \sum_{i=1}^{n} d_i v_i - \sum_{i=1}^{r} t_i p_i = 0$$

$$\rightarrow \sum_{i=1}^{r} (c_i - t_i) p_i + \sum_{i=1}^{n} d_i v_i = 0$$

و $\{p_1, \cdots, p_r, v_1, \cdots, v_n\}$ پایه‌ای برای W_1 است که این مجموعه مستقل خطی است؛ در نتیجه، $d_1 = \cdots = d_n = 0$ و با جای‌گزینی در * داریم:

$$\sum_{i=1}^{r} c_i p_i + 0 + \sum_{i=1}^{m} d'_i w_i = 0$$

$$\rightarrow \sum_{i=1}^{r} c_i p_i + \sum_{i=1}^{m} d'_i w_i = 0$$

و $\{p_1, \cdots, p_r, w_1, \cdots, w_n\}$ پایه‌ای برای W_2 است که این مجموعه مستقل خطی است؛ در نتیجه، $c_1 = \cdots = c_r = 0$ و $d'_1 = \cdots = d'_m = 0$، پس $\{p_1, \cdots, p_r, v_1, \cdots, v_n, w_1, \cdots, w_n\}$ مجموعه‌ای مستقل خطی برای $W_1 + W_2$ است و هم‌چنین،

$$r+n+m = dim(W_1+W_2) = dim(W_1)+dim(W_2)-dim(W_1 \cap W_2) = (r+n)+(r+m)-r$$

مثال: فرض کنید $V = span(\{v_1, \cdots, v_6\})$ به طوری که:

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} -1 \\ 2 \\ -2 \\ 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 \\ -3 \\ 0 \\ 3 \end{bmatrix}, \quad v_5 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad v_6 = \begin{bmatrix} 0 \\ 0 \\ 2 \\ -2 \end{bmatrix}.$$

پایه‌ای برای فضای خطی V بیابید.

حل: فرض کنید که v_i ستون iام ماتریس A باشد، $A = \begin{bmatrix} v_1 & \cdots & v_6 \end{bmatrix}$. در این صورت $V = C(A)$. بنابراین برای یافتن پایه‌ای برای فضای V باید ابتدا بعد $C(A)$ و سپس بردارهای مستقلی که آن را تولید می‌کنند بیابیم. ماتریس تحویل یافته‌ی سطری پلکانی A برابر است با:

$$R = \begin{bmatrix} 1 & 0 & -3 & 0 & 0 & 4 \\ 0 & 1 & 1 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

ادعا می‌کنیم که v_1, v_2, v_4, v_5 بردارهای مستقل خطی بوده و بردارهای v_3 و v_6 وابسته‌ی خطی هستند؛ به عبارت دیگر، بردارهای مستقل خطی، همان بردارهای متناظر با بردارهای محوری هستند. می‌دانیم که ماتریس وارون‌پذیر E وجود دارد که $EA = R$. ستون‌های ماتریس R را با $R_1, \cdots, R_6$ نمایش می‌دهیم. واضح است که $R_3 = -3R_1 + R_2$ و $R_6 = 4R_1 - 2R_2 - \frac{2}{3}R_3$ و R_1, R_2, R_4, R_5 ستون‌های مستقل R هستند. می‌توان به راحتی دید که $v_3 = -3v_1 + v_2$ و $v_6 = 4v_1 - 2v_2 - \frac{2}{3}v_3$. همچنین فرض کنید $c_1v_1 + c_2v_2 + c_3v_4 + c_4v_5 = 0$. چون $EA = R$ داریم:

$$E(c_1v_1 + c_2v_2 + c_3v_4 + c_4v_5) = c_1Ev_1 + c_2Ev_2 + c_3Ev_4 + c_4Ev_5$$

$$= c_1R_1 +_2 R_2 + c_3R_4 + c_4R_5 = 0$$

چون R_1, R_2, R_4, R_5 مستقل خطی هستند، پس $c_1 = c_2 = c_3 = c_4 = 0$ که نشان می‌دهد ادعای استقلال خطی v_1, v_2, v_4, v_5 درست است. پس $B = \{v_1, v_2, v_4, v_5\}$ پایه‌ای برای V است.

گزاره: فرض کنید $A \in M_{mn}(R)$ و R ماتریس تحویل یافته‌ی سطری پلکانی آن باشد. در این صورت:
$$dim\, C(A) = dim\, C(R).$$

برهان: ستون‌های R را با $R_1, \cdots , R_n$ نمایش می‌دهیم. ستون‌های محوری، مستقل خطی هستند. فرض کنید $R_{i1}, \cdots , R_{ir}$ ستون‌های محوری باشند؛ در این صورت $C(R) = span(\{R_{i_1}, \cdots , R_{i_r}\})$.
ثابت می‌کنیم که $\{A_{i_1}, \cdots , A_{i_r}\}$ ــ که در آن، A_i ستون i ام A است ــ نیز مستقل‌خطی هستند.
فرض کنید $c_1 A_{i_1} + \cdots + c_r A_{i_r} = 0$. می‌دانیم که ماتریس وارون‌پذیر E وجود دارد که $EA = R$. بنابراین:

$$E(c_1 A_{i_1} + \cdots + c_r A_{i_r}) = c_1 EA_{i_1} + \cdots + c_r EA_{i_r} = c_1 R_{i_1} + \cdots + c_r R_{i_r} = 0$$

بنابراین، استقلال خطی $R_{i_1}, \cdots , R_{i_r}$ نشان می‌دهد که $c_1 = \cdots = c_r = 0$ و در نتیجه، $r \leq dim\, C(A)$.
حال با فرایند مشابه نشان می‌دهیم که اگر ستون‌های $A_{i_1}, \cdots , A_{i_m}$ پایه‌ای برای فضای ستونی A باشند (یعنی $dim\ C(A) = m$)، آنگاه $R_{i_1}, \cdots , R_{i_m}$ نیز مستقل خطی هستند. چون $dim\, C(R) = r$ بنابراین $m \leq r$ و در نتیجه $dim\, C(R) = r$.
برای اثبات، چون می‌دانیم $A = E^{-1} R$، اگر $c_1 R_{i_1} + \cdots + c_m R_{i_m} = 0$، آنگاه $c_1 A_{i_1} + \cdots + c_m A_{i_m} = 0$ و در نتیجه $c_1 = \cdots = c_m = 0$.
گزاره: اگر $A \in M_{mn}(R)$ و $P \in M_{mn}(R)$ وارون‌پذیر باشند، آنگاه:

$$dim\, C(PA) = dim\, C(A).$$

برهان: مشابه آن‌چه دیدیم، اگر در نظر بگیریم $S = PA$، آنگاه چون P وارون‌پذیر است، $A = P^{-1} S$ و در نتیجه اگر $c_1 S_{i_1} + c_2 S_{i_2} + \cdots + c_m S_{i_m} = 0$، داریم $c_1 A_{i_1} + \cdots + c_m A_{i_m} = 0$ و در نتیجه چون PA برابر S بود، $c_1 = \cdots = c_m = 0$.

نکته: اگر $A \in M_{mn}(R)$ و $P_{mn}(R)$ وارون‌پذیر باشند، ممکن است $C(PA)$ و $C(A)$ برابر نباشند؛ در حالی که بنابر گزاره‌ی پیشین، $dim\, C(A) = dim\, C(PA)$.

مثال: اگر $A = \begin{bmatrix} ۱ & ۱ \\ ۰ & ۰ \end{bmatrix}$ و $P = \begin{bmatrix} ۰ & ۱ \\ ۱ & ۰ \end{bmatrix}$ آن‌گاه داریم:

$$C(A) = span(\{\begin{bmatrix} ۱ \\ ۰ \end{bmatrix}\}) \quad , \quad C(PA) = span(\{\begin{bmatrix} ۰ \\ ۱ \end{bmatrix}\})$$

سوال: اگر $A \in M_{mn}(R)$ و ماتریس $P \in M_m(R)$ وارون‌پذیر باشد، فضای تولید شده توسط سطرهای A و سطرهای PA چه رابطه‌ای با هم دارند؟

تعریف: فرض کنید $A \in M_{mn}(R)$. فضای تولیدشده توسط سطرهای A را **فضای سطری** A می‌نامند.

گزاره: اگر $A \in M_{mn}(R)$ و ماتریس $P \in M_m(R)$ وارون‌پذیر باشد، آنگاه فضای سطری A و فضای سطری PA یکسان هستند.

برهان: اگر A_i سطر iام ماتریس A باشد، داریم:

$$PA = \begin{bmatrix} p_{۱۱} & \cdots & p_{۱m} \\ \vdots & & \vdots \\ p_{m۱} & \cdots & p_{mm} \end{bmatrix} \begin{bmatrix} A_۱ \\ \vdots \\ A_m \end{bmatrix} = \begin{bmatrix} p_{۱۱}A_۱ + \cdots + p_{۱m}A_m \\ \vdots \\ p_{m۱}A_۱ + \cdots + p_{mm}A_m \end{bmatrix}$$

$$\Rightarrow \quad PA \text{ فضای سطری } = span(\{p_{i۱}A_۱ + \cdots + p_{im}A_m\}_{۱\le i\le m})$$

$$\subseteq span(\{A_۱, \cdots, A_m\})$$

پس فضای سطری PA زیرمجموعه مساوی فضای سطری A است. چون P وارون‌پذیر است، با استدلال مشابه فضای سطری A نیز زیرمجموعه مساوی فضای سطری PA است و در نتیجه حکم ثابت می‌شود.

قضیه: اگر $A \in M_{mn}(R)$، آنگاه $dimN(A) + dimC(A) = n$.

برهان:

اگر R ماتریس تحویل شده‌ی سطری پلکانی A باشد، چون ماتریس وارون‌پذیر E وجود دارد که $EA = R$، در نتیجه $dimC(R) = dimC(A)$ و هم‌چنین $N(A) = N(R)$. ستون‌های محوری R پایه‌ای برای فضای $C(R)$ هستند و از طرفی معادله‌ی $Ax = ۰$ شامل $n - dimC(A)$ متغیر آزاد و $dimC(A)$ متغیر محوری است، در نتیجه حکم ثابت می‌شود.

تعریف:

فضای پوچ چپ A را فضای پوچ A^T تعریف می‌کنیم.

چهار زیرفضای بنیادی

به ازای ماتریس $A \in M_{mn}(R)$، فضای ستونی A یا همان $C(A)$، فضای پوچ A یا همان $N(A)$، فضای سطری A یا همان $C(A^T)$ و فضای پوچ چپ A یا همان $N(A^T)$ را چهار زیرفضای بنیادی A می‌نامند. بنابر قضیه:

$$dimC(A) + dimN(A) = n$$

$$dimC(A^T) + dimN(A^T) = m$$

قضیه: بعد فضای سطری و فضای ستونی ماتریس $A \in M_{mn}(R)$ با هم یکسان است.

برهان: فرض کنید R ماتریس تحویل شده‌ی سطری پلکانی A باشد؛ در این صورت، تعداد ستون‌های محوری آن برابر تعداد سطرهای ناصفر R و برابر با رتبه‌ی A است؛ از طرفی،

$$dimC(A) = dimC(R) = rank(A)$$

$$dimC(A^T) = rank(A)$$

و بنابراین، حکم ثابت می‌شود.

۴

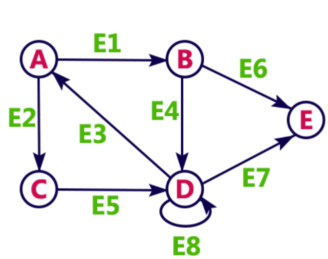
جبر خطی و گراف‌ها

یادآوری: در این جلسه (جلسه‌ی نهم)، زیرفضاهای بنیادی ماتریس A یادآوری شد، دو مثال از آن در اسلایدهای ۵ و ۶ حل شد و سپس بخش مربوط به گراف‌ها و شبکه‌ها توضیح داده شد.

گراف‌ها

هر گراف با زوج مرتب (V, E) معرفی می‌شود که در آن V مجموعه‌ی رئوس گراف و E مجموعه‌ی یال‌های گراف است. فرض کنید $|V| = n$ و $|E| = m$ باشد. در این صورت ماتریس وقوع یال ـ رأس $A_{m\times n}$ را به این صورت به آن نسبت می‌دهیم ($A \in M_{mn}(R)$): اگر یالی از رأس j به رأس k باشد، در سطر متناظر آن یال در ستون jام ۱ـ گذاشته و در ستون kام ۱+ می‌گذاریم؛ در غیر این صورت، در این درایه‌ها، صفر قرار می‌دهیم.

مثال: در زیر، ماتریس A^T مربوط به گراف نمایش داده شده، نوشته شده است:



$$
\begin{array}{c}
\\ A \\ B \\ C \\ D \\ E
\end{array}
\begin{array}{c}
\begin{array}{cccccccc} E1 & E2 & E3 & E4 & E5 & E6 & E7 & E8 \end{array} \\
\begin{pmatrix}
1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & -1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & -1 & -1 & 0
\end{pmatrix}
\end{array}
$$

A^T را ماتریس وقوع رأس ـ یال می‌نامند.(در نظریه‌ی گراف معمولاً A^T را ماتریس وقوع می‌نامند).

حال، چهار زیرفضای بنیادی مربوط به گراف A را بررسی می‌کنیم و معادلی که در گراف دارند را

تشریح می‌کنیم. همچنین به اختلاف پتانسیل و شدت جریان در مدار الکتریکی که قابل بازنویسی با یک گراف (جهت‌دار) می‌باشند می‌پردازیم.

فضای پوچ A

در نگاه اول در می‌یابید که حاصل‌جمع ستون‌ها صفر است؛ بنابراین اگر A_i ستون iام A باشد، داریم:

$$A_۱ + \cdots + A_n = \begin{bmatrix} A_۱ & \cdots & A_n \end{bmatrix} \begin{bmatrix} ۱ \\ \vdots \\ ۱ \end{bmatrix} = A \begin{bmatrix} ۱ \\ \vdots \\ ۱ \end{bmatrix} = ۰ \quad \Rightarrow \quad \begin{bmatrix} ۱ \\ \vdots \\ ۱ \end{bmatrix} \in N(A)$$

از طرفی از حل $Ax = ۰$، با توجه به نحوه‌ی قرارگیری ۱ و ۱− در سطر اول، به $x_۱ - x_۲ = ۰$ می‌رسیم. اگر $x_۱ - x_۲$ را پتانسیل هر کدام از رئوس فرض کنید، اختلاف پتانسیل دو سر یال ۱ برابر است با $x_۱ - x_۲$. به طور مشابه به ازای هر سطر (که معادل یال است).

چون گراف هم‌بند است، پس:

$$x_۱ = x_۲ = \cdots = x_n$$

بنابراین:

$$x = c \begin{bmatrix} ۱ \\ \vdots \\ ۱ \end{bmatrix}$$

پس $N(A)$ با بردار $\begin{bmatrix} ۱ \\ \vdots \\ ۱ \end{bmatrix}$ تولید می‌شود و $dim\, N(A) = ۱$. بنابر قضیه‌ای که دیدیم :

$$dim\, N(A) + dim\, C(A) = n$$

پس:

$$dim\, C(A) = ۴ - ۱ = ۳$$

فضای ستونی

دیدیم که $dim\, C(A) = ۳$. بنابراین بعد فضای خطی همه‌ی بردارهای b که $Ax = b$، برابر با ۳

است. همان‌طور که دیدید از حل $Ax = b$ به معادلات $b_۱ + b_۳ = b_۲$ و $b_۳ + b_۵ = b_۴$ می‌رسیم.
پس b در فضای ستونی A است اگر و تنها اگر در معادله‌ی فوق صدق کند.
از حل معادله به دست می‌آید که:

$$x_۳ - x_۲ = b_۳ \quad , \quad x_۳ - x_۱ = b_۲ \quad , \quad x_۲ - x_۱ = b_۱$$

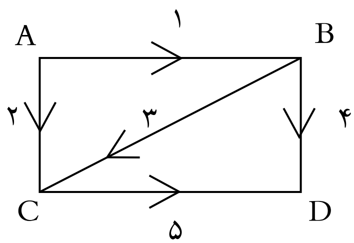
$$\underbrace{(x_۲ - x_۱) + (x_۳ - x_۲) = x_۳ - x_۱}_{\text{مجموع اختلاف پتانسیل در یک حلقه صفر است.(KVL)}}$$

$$b_۱ + b_۳ = b_۲$$

به طریق مشابه $b_۳ + b_۵ = b_۴$ که تعبیر آن KVL در حلقه‌ی دوم است.

فضای پوچ چپ

از تلاش برای حل معادله‌ی $A^T y = ۰$ به‌دست می‌آوریم (راس‌ها و یال‌ها، نام‌گذاری شده‌اند):



راس A: $-y_۱ - y_۲ = ۰$

راس B: $y_۱ - y_۳ - y_۴ = ۰$

راس C: $y_۲ + y_۳ - y_۵ = ۰$

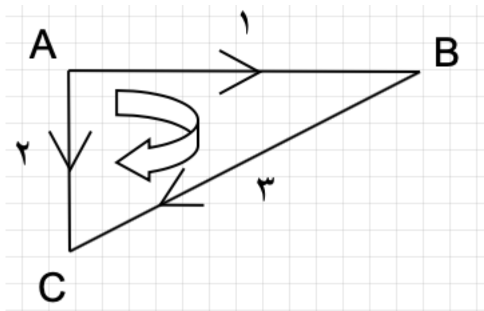
راس D: $y_۴ + y_۵ = ۰$

اگر y_i ها را جریان عبوری فرض کنیم، معادلات فوق، مجموع جریان هر راس را نشان می‌دهند.

۱ ــ تلاش می‌کنیم برای فضای خطی yهایی که $A^T y = ۰$ (معادلاً فضای پوچ چپ A)، یک پایه بیابیم.
ابتدا به این سوال پاسخ می‌دهیم که بعد این فضا چیست؟ بعد A برابر با ۳ بود؛ در نتیجه:

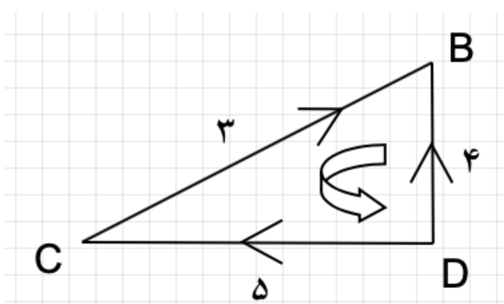
$$dim\, N(A^T) = m - r = ۵ - ۳ = ۲$$

پس بایستی برداری مستقل خطی بیابیم که $A^Ty = ۰$. از معادلات (*) دریافتیم که جریان در هیچ رأسی ذخیره نمی‌شود و در همه‌ی گراف از جمله داخل حلقه‌های کوچک می‌چرخد. حال به حلقه‌ها نگاه می‌کنیم.



$$y = \begin{bmatrix} ۱ \\ -۱ \\ ۱ \\ ۰ \\ ۰ \end{bmatrix}$$

$$\begin{bmatrix} -۱ & -۱ & ۰ & ۰ & ۰ \\ ۱ & ۰ & -۱ & -۱ & ۰ \\ ۰ & ۱ & ۱ & ۰ & -۱ \\ ۰ & ۰ & ۰ & ۱ & ۱ \end{bmatrix} \begin{bmatrix} ۱ \\ -۱ \\ ۱ \\ ۰ \\ ۰ \end{bmatrix} = \begin{bmatrix} -۱ + ۱ \\ ۱ + ۰ - ۱ \\ -۱ + ۱ + ۰ \\ ۰ \end{bmatrix} = \begin{bmatrix} ۰ \\ ۰ \\ ۰ \\ ۰ \end{bmatrix}$$



$$y = \begin{bmatrix} ۰ \\ ۰ \\ ۱ \\ -۱ \\ ۱ \end{bmatrix}$$

$$A^T y' = ۰$$

در نتیجه، دو بردار فوق در $N(A^T)$ قرار داشته، مستقل خطی بوده و $dim(N(A^T)) = ۲$، پس پایه‌ای برای $N(A^T)$ هستند. در غیاب جریان‌های منبع، هر بردار در $N(A^T)$ نشانگر یک بردار جریان در مدار است که در حال چرخش است، یعنی KCL (مجموع جریان‌ها در هر راس صفر است). حال اگر در رئوسی از مدار، مشابه مثال مدار ارائه شده، منبع جریان داشته باشیم، برای محاسبه‌ی جریان، باید به محاسبه‌ی $A^T y = f$ بپردازیم که در حضور منابعی از جریان $A^T y = f$ در رأس‌ها متناظر قانون KCL است که کل جریان ورودی برابر با جریان خروجی است.
بنابراین فضاهای بنیادی منسوب به A را مطالعه کردیم. حال فرض کنید مدار داده شده است (مشابه اسلایدها)؛ می‌توان برای محاسبه‌ی y (جریان‌ها) از معادله‌ی $A^T = f$ و برای محاسبه‌ی پتانسیل‌ها از $A^T cAx = A^T cb - f$، مطابق محاسبات انجام شده در اسلایدها، استفاده کرد.

۵

جلسه دهم

می‌دانیم فضای ستونی A که با $C(A)$ نشان داده می‌شود، فضای تولیدشده توسط ستون‌های A ــ که $A_1, \cdots, A_n$ هستند ــ است:

$$C(A) = span(\{A_1, \cdots, A_n\}) = \{x_1 A_1 + \cdots + x_n A_n | x \in R^n\} .$$

به تعبیر دستگاه معادلات، $C(A) \subseteq R^m$ شامل بردارهای $b \in R^m$ است که به ازای آن‌ها، معادله‌ی $Ax = b$ جواب داشته باشد؛ به عبارت دیگر، $x \in R^n$ وجود داشته باشد به طوری که $Ax = x_1 A_1 + \cdots + x_n A_n = b$.

شرط لازم و کافی برای وجود وارون راست ماتریس $A \in M_{mn}(R)$
فرض کنید $c \in Mnm(R)$ وجود دارد به طوری که $AC = I_m$. اگر ستون iام ماتریس C را با C_i نمایش دهیم،

$$AC = A\begin{bmatrix} C_1 & \cdots & C_m \end{bmatrix} = \begin{bmatrix} AC_1 & \cdots & AC_m \end{bmatrix} = \begin{bmatrix} e_1 & \cdots & e_m \end{bmatrix}$$

$$\Rightarrow \forall i \ 1 \leq i \leq m \ : AC_i = e_i \Rightarrow \forall b \in R^m \ : \ b = b_1(AC_1) + \cdots + b_m(AC_m)$$

$$\Rightarrow \forall b \in R^m \ : \ b = A(b_1 C_1 + \cdots + b_m C_m) .$$

به بیان دیگر، $b_1 C_1 + \cdots + b_m C_m$ جواب معادله‌ی $Ax = b$ است.

$$C(A) = R^m \Rightarrow dim\, C(A) = rank\, A = m .$$

پس همه‌ی m سطر ماتریس A مستقل خطی هستند. به عبارتی، ماتریس A دارای «رتبه‌ی سطری کامل» [1] است.

حال به بررسی عکس ماجرا می‌پردازیم و نشان می‌دهیم اگر m سطر ماتریس مستقل خطی باشند، وارون راست وجود دارد؛ که به نوعی با برعکس کردن فلش‌های استدلال فوق قابل انجام است:

فرض کنید m سطر A مستقل خطی باشد؛ بنابراین $C(A) = R^m$. پس معادله‌ی $Ax = e_i$ به ازای هر $1 \leq i \leq m$ جواب دارد.

جواب معادله‌ی $Ax = e_i$ را با C_i نشان دهید و به عنوان ستون iام ماتریس C در نظر بگیرید. به وضوح $AC = I_m$.

نکته: ماتریس $A \in M_{mn}(R)$ وارون راست دارد اگر و تنها اگر m سطر A مستقل خطی باشند.

شرط لازم و کافی برای وجود وارون چپ ماتریس $A \in M_{mn}(R)$

فرض کنید $B \in M_{nm}(R)$ به طوری که $BA = I_n$. سطر i ام ماتریس B را با B_i نمایش می‌دهیم. لذا:

$$BA = \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix} A = \begin{bmatrix} B_1 A \\ \vdots \\ B_n A \end{bmatrix} = \begin{bmatrix} e_1^T \\ \vdots \\ e_n^T \end{bmatrix} \Rightarrow \forall\ 1 \leq i \leq n\ \ \underbrace{B_i = e_i^T}_{A^T B_i^T = e_i}$$

$$\Rightarrow \forall x \in R^n\ :\ x = x_1 A^T B_1^T + \cdots + x_n A^T B_n^T = A^T(x_1 B_1^T + \cdots + x_n B_n^T)\ .$$

پس به ازای هر $x \in R^n$، معادله‌ی $A^T b = x$ جواب دارد و جواب آن $b = x_1 B_1^T + \cdots + x_n B_n^T$ است، بنابراین $C(A^T) = R^n$ و در نتیجه:

$$rank(A) = dim\, C(A^T) = n \quad .$$

پس ستون‌های A مستقل خطی هستند.

حال اگر ستون‌های A مستقل خطی باشند (یا به عبارتی، A یک ماتریس با «رتبه‌ی ستونی کامل» [2] باشد)، A وارون چپ دارد. به طریق مشابه، کم و بیش با برعکس کردن فلش‌های استدلال فوق به جواب می‌رسیم:

می‌دانیم $C(A^T) = R^n$. در نتیجه معادلات $A^T b = e_i$ به ازای هر $1 \leq i \leq n$ جواب دارد. ستون B_i ماتریس B را جواب معادله‌ی $A^T b = e_i$ در نظر بگیرید، آنگاه $BA = I_n$.

---

[1] full row rank [2] full column rank

نکته: ماتریس $A \in M_{mn}(R)$ وارون چپ دارد اگر و تنها اگر n ستون A مستقل خطی باشند.

توجه: در فصل ۳ ثابت می‌کنیم که:
۱ـ A^TA وارون‌پذیر است اگر رتبه‌ی A، n باشد.
۲ـ AA^T وارون‌پذیر است، اگر رتبه‌ی A، m باشد.

اگر وارون راست (چپ) ماتریس A وجود داشته باشد، می‌توان مشابه مثالی که در اسلایدها موجود است، به شرط دانستن نکته‌ی فوق، وارون راست (چپ) A را برحسب A محاسبه نمود.

توجه: در اسلاید ۱۴ اشاره شده است که فضای پوچ ماتریس وقوع یال ـ رأس $(edge - node\ incident\ matrix)$ توسط بردار $\begin{bmatrix} ۱ \\ \vdots \\ ۱ \end{bmatrix}$ تولید می‌شود. در ادامه، این نکته را به ازای هر گراف جهت‌دار هم‌بند $(connected\ directed\ graph)$ دل‌خواه نشان می‌دهیم.

فرض کنید $G = (V, E)$ که در آن $V = \{v_۱, \cdots, v_n\}$ باشد؛ در این صورت، هر یال، یک راس خروجی و یک راس ورودی دارد، پس در سطر i ام، یک درایه‌ی منفی یک مربوط به ستون j ام و یک درایه‌ی یک مربوط به ستون k ام وجود دارد و بقیه صفر هستند؛ بنابراین در معادله‌ی $Ax = ۰$، داریم $-x_j + x_k = ۰$ ظاهر می‌شود. چون گراف هم‌بند است، پس بین هر دو راس، یک مسیر وجود دارد، پس متغیر مربوط به هر دو راس با هم برابرند و در نتیجه $x_۱ = \cdots = x_n = c$، پس:

$$N(A) = span(\{\begin{bmatrix} ۱ \\ \vdots \\ ۱ \end{bmatrix}\}) .$$

فرض کنید $A \in M_{mn}(R)$. فضای پوچ A، زیرفضای خطی از R^n است که تحت اثر A روی آن‌ها به مجموعه‌ی تک عضوی بردار صفر می‌رود. چون Ax ترکیبی از ستون‌های A است، هر برداری (تحت ضرب A) به فضای ستونی می‌رود؛ بنابراین عملکرد A را می‌توان به صورت یک تابع $T : R^n \to R^m$ دید که در آن $T(x) = Ax$.
این تابع دو ویژگی مهم دارد. به ازای هر دو بردار $x, y \in R^n$ و $c \in R$ داریم:

$$T(x + y) = A(x + y) = Ax + Ay = T(x) + T(y)$$

و

$$T(cx) = A(cx) = cAx = cT(x)\ .$$

بنابراین تابع T نسبت به عملگر جمع و ضرب اسکالر خطی عمل می‌کند. لذا به طور کلی، تابع خطی روی فضاهای خطی را چنین تعریف می‌کنیم:

فرض کنید w, v دو فضای خطی باشند. نگاشت $f : v \to w$ را **تابع خطی (تبدیل خطی)** گویند هرگاه به ازای هر $x, y \in V$ و $c \in R$:

$$T(cx) = cT(x)$$

$$T(x+y) = T(x) + T(y)$$

نکته: به سادگی می‌توان دید $T(۰) = ۰$:

$$T(۰) = T(۰+۰) = ۲T(۰) \ \Rightarrow\ T(۰) = ۲T(۰) \ \Rightarrow\ T(۰) = ۰$$

فضای پوچ و برد توابع خطی

اگر $T : V \to W$ یک تابع خطی باشد، آن‌گاه:

$$N(T) := \{x \in v | T(x) = ۰\} \quad , \quad Im(T) = \{T(x) | x \in v\}\ .$$

رتبه‌ی T را نیز بعد فضای $Im(T)$ تعریف می‌کنیم و قرار می‌دهیم $rank(T) = dim(Im(T))$ (توجه کنید که به راحتی می‌توان دید که $N(T)$ و $Im(T)$ زیرفضای خطی هستند).

توابع یک‌به‌یک (injective) و پوشا (surjective)

تعریف: فرض کنید $T : V \to W$ یک تابع خطی باشد. اگر $Im(T) = w$، آن‌گاه T یک تابع خطی پوشاست.

گزاره: تابع خطی T یک‌به‌یک است اگر و تنها اگر $N(T) = \{۰\}$.

برهان: (خلف) ابتدا فرض کنید T یک‌به‌یک باشد و $x \in N(T)$ و $x \neq ۰$؛ بنابراین $T(۰) = ۰ = T(x)$ که تناقض است. در نتیجه، $N(T) = ۰$.

حال فرض کنید $N(T) = \{۰\}$ و $T(x) = T(y)$ و $x \neq y$ (فرض خلف). در این صورت:

$$T(x) - T(y) = T(x-y) = ۰ \quad \Rightarrow \quad ۰ \neq x - y \in N(T)$$

پس $N(T) = 0$ و به تناقض می‌رسیم؛ پس T یک‌به‌یک است.

نکته: فرض کنید V یک فضای برداری با بعد متناهی و پایه‌ی $\{v_1, \cdots, v_n\}$، و W یک فضای خطی برداری باشد به طوری که $\{w_1, \cdots, w_n\} \subseteq W$؛ در این صورت تابع خطی یکتای $T : V \to W$ وجود دارد که $T(v_i) = w_i$ به ازای هر $1 \leq i \leq n$.

برهان: با داشتن مقدار T روی عناصر پایه‌ی V، مقدار T را روی همه‌ی اعضای V داریم، زیرا هر بردار در $v \in V$ را می‌توان به صورت ترکیب خطی اعضای پایه نوشت، یعنی $V = c_1v_1 + \cdots + c_nv_n$؛ در نتیجه، چون T تابعی خطی است، داریم:

$$T(v) = T(c_1v_1 + \cdots + c_nv_n) = c_1T(v_1) + \cdots + c_nT(v_n) = c_1w_1 + \cdots + c_nw_n \quad .$$

واضح است که این تابع خطی به طور یکتا تعیین می‌شود، زیرا به ازای هر $v \in V$، $c_1, \cdots, c_n$ به طور یکتا وجود دارد که $v = \sum_{i=1}^{n} c_iv_i$ و هم‌چنین $T(v) = \sum_{i=1}^{n} c_iT(v_i)$؛ در نتیجه، مقدار $T(v)$ با مقادیر $T(v_i)$ مشخص می‌گردد.

نکته: اگر $T : V \to W$ یک تابع خطی باشد، $dim(Im(T)) + dim(N(T)) = n$ که در آن $n = dim(V)$.

برهان (اختیاری):

ایده‌ی اثبات، در نظر گرفتن یک پایه برای زیرفضای خطی از V و گسترش آن به پایه‌ای برای V است که قبلا به عنوان رویه‌ای برای ساختن پایه برای فضا مطرح کردیم. فرض کنید $dim(N(T)) = k$؛ در نتیجه، k بردار مستقل خطی $v_1, \cdots, v_k$ وجود دارد که زیرفضای خطی $N(T)$ را تولید کند، در نتیجه برای هر i از صفر تا k داریم $T(v_i) = 0$. مجموعه‌ی مستقل خطی $\{v_1, \cdots, v_k\}$ را به یک پایه برای V گسترش می‌دهیم، پس بردارهای $\{v_{k+1}, \cdots, v_n\}$ وجود دارند که $\{v_1, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ پایه‌ای برای V است. ثابت می‌کنیم که $\{T(v_{k+1}), \cdots, T(v_n)\}$ پایه‌ای برای $Im(T)$ است، یعنی هم استقلال در آن برقرار است و هم اعضای فضا را تولید می‌کند.

برای اثبات استقلال، فرض کنید $c_{k+1}T(v_{k+1}) + \cdots + c_nT(v_n) = 0$؛ نشان می‌دهیم این c_i ها صفرند:

$$c_{k+1}T(v_{k+1}) + \cdots + c_nT(v_n) = T(c_{k+1}v_{k+1} + \cdots + c_nv_n) = 0$$

$$\to c_{k+1}v_{k+1} + \cdots + c_nv_n \in N(T)$$

و چون مجموعه‌ی v_i ها به ازای i از ۱ تا k پایه‌ای برای $N(T)$ است، پس بردارهای d_1 تا d_k در R وجود دارند، به گونه‌ای که:

$$c_{k+1}v_{k+1} + \cdots + c_n v_n = d_1 v_1 + \cdots + d_k v_k$$

(توجه کنید که $\sum c_i v_i$ به عنوان عضوی از $N(T)$، ترکیبی خطی از اعضای پایه‌ی $N(T)$ است)

$$\to d_1 v_1 + \cdots + d_k v_k - c_{k+1}v_{k+1} + \cdots + c_n v_n = 0$$

چون چون مجموعه‌ی v_i ها به ازای i از ۱ تا n پایه است، پس مستقل خطی است و در نتیجه:

$$d_1 = \cdots = d_k = 0$$

$$c_{k+1} = \cdots = c_n = 0$$

در نتیجه، مجموعه‌ی $\{T(v_{k+1}), \cdots, T(v_n)\}$ مستقل خطی است.

حال ثابت می‌کنیم که این مجموعه، تولیدکننده‌ی اعضای (مولد) $Im(T)$ نیز است.

فرض کنید w در $Im(T)$ باشد؛ در نتیجه، بردار v در V وجود دارد به طوری که $w = T(v)$. از طرفی، چون بردار v در V است، پس ترکیبی خطی از اعضای پایه‌ی V است، یعنی:

$$v = c_1 v_1 + \cdots + c_k v_k + c_{k+1}v_{k+1} + \cdots + c_n v_n$$

$$\to T(v) = c_1 T(v_1) + \cdots + c_k T(v_k) + c_{k+1}T(v_{k+1}) + \cdots + c_n T(v_n)$$

و چون مجموعه‌ی v_i ها به ازای i از ۱ تا k پایه‌ای برای $N(T)$ است، پس $T(v_i)$ ها به ازای i از ۱ تا k صفر هستند:

$$\to T(v) = w = c_{k+1}T(v_{k+1}) + \cdots + c_n T(v_n)$$

پس مجموعه‌ی $\{T(v_{k+1}), \cdots, T(v_n)\}$ یک مولد برای $Im(T)$ بوده و چون استقلال نیز داشت، پس یک پایه برای $Im(T)$ است.

۶

ضرب داخلی

مثال: فرض کنید V فضای خطی شامل همه‌ی توابع پیوسته حقیقی روی بازه‌ی $[۰, ۲\pi]$ است. ضرب داخلی روی V به این صورت تعریف می‌شود که به ازای هر تابع $f, g : [۰, ۲\pi] \to R$:

$$< f, g >= \int_۰^{۲\pi} f(x)g(x)dx$$

بنابراین، نرم هر تابع $f : [۰, ۲\pi] \to R$ برابر است با:

$$\| f \|= \int_۰^{۲\pi} f(x)^۲ dx$$

پرسش: نشان دهید $\{\frac{۱}{\sqrt{\pi}} sin(x), \frac{۱}{\sqrt{\pi}} cos(x)\}$ مجموعه‌ای متعامد و یکه است.
حل: به راحتی می‌توان محاسبه کرد که:

$$\| sin\ x \|^۲= \int_۰^{۲\pi} sin^۲ x dx = \frac{۱}{۲}(۲\pi - \int_۰^{۲\pi} cos۲x) = \pi$$

$$\| cos\ x \|^۲= \int_۰^{۲\pi} cos^۲ x dx = \frac{۱}{۲}(۲\pi - \int_۰^{۲\pi} sin۲x) = \pi$$

$$< sinx, cosx >= \int_۰^{۲\pi} sinxcosxdx = \frac{۱}{۲} \int_۰^{۲\pi} sin۲xdx = ۰$$

مثال: فرض کنید V فضای خطی همه‌ی چندجمله‌ای‌های حداکثر درجه n، $n \geq ۲$، روی $[-۱, ۱]$ است. ضرب داخلی روی V را مطابق اولین رابطه‌ی این صفحه در نظر بگیرید. آیا مجموعه‌ی

$\{1, x, x^2\}$ در V متعامد است؟ در صورتی که پاسخ خیر است، با استفاده از فرایند گرام اشمیت پایه‌ای متعامد یکه برای زیرفضای $W = span(\{1, x, x^2\})$ ، زیرفضای چندجمله‌ای های حداکثر درجه ۲ به دست آورید.

حل:

$$< 1, x >= \int_{-1}^{1} x dx = \frac{x^2}{2}|_{-1}^{1} = 0$$

بنابراین عناصر ۱ و x متعامدند.

$$< x, x^2 >= \int_{-1}^{1} x^3 dx = \frac{x^4}{4}|_{-1}^{1} = 0$$

عناصر x و x^2 نیز بر هم عمودند.

$$< 1, x^2 >= \int_{-1}^{1} x^2 dx = \frac{x^3}{3}|_{-1}^{1} = \frac{2}{3} \neq 0$$

بنابراین مجموعه‌ی $\{1, x, x^2\}$ متعامد نیست. قرار دهید:

$$q_1 = \frac{1}{\| 1 \|} \quad , \quad q_2 = \frac{x}{\| x \|}$$

بنابراین داریم:

$$Q_3 = x^2 - < x^2, q_1 > q_1 - < x^2, q_2 > q_2 = x^2 - < x^2, q_1 > q_1$$

$$= x^2 - \frac{< x^2, 1 >}{\sqrt{< 1, 1 >}} \frac{1}{\sqrt{< 1, 1 >}} = x^2 - \frac{< x^2, 1 >}{< 1, 1 >} 1$$

$$= x^2 - \frac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} 1 dx} = x^2 - \frac{\frac{x^3}{3}|_{-1}^{1}}{x|_{-1}^{1}} = x^2 - \frac{\frac{1}{3}(1+1)}{1+1} = x^2 - \frac{1}{3} .$$

مجموعه‌ی $\{1, x, x^2 - \frac{1}{3}\}$ مجموعه‌ی متعامد و یکه، حاصل اعمال فرآیند گرام‌ـاشمیت بر روی $\{0, x, x^2\}$ است.

مثال: بهترین تقریب $y = x^5$ در بازه‌ی $0 \leq x \leq 1$ با خط راست $y = Dx + C$ را بیابید.

حل: فرض کنید $V = P_n(x)$ فضای همه‌ی چندجمله‌ای‌های حداکثر درجه n، $n \geq 5$، با ضرب

داخلی معرفی‌شده باشد. با درنظرگیری $W = span(\{۱, x\})$، مسئله، یافتن بهترین تقریب برای $x^۵$ توسط بردارهای W است. مجموعه‌ی $\{۱, x\}$ روی بازه‌ی $[۰, ۱]$ متعامد نیست، زیرا:

$$< ۱, x >= \int_۰^۱ x dx = \frac{x^۲}{۲}|_۰^۱ = \frac{۱}{۲} \neq ۰$$

بنابراین ابتدا یک پایه‌ی متعامد یکه برای W با استفاده از فرایند گرام‌ـ‌اشمیت به دست می‌آوریم:

$$q_۱ = \frac{۱}{\sqrt{< ۱, ۱ >}} .$$

$$Q_۲ = x- < x, q_۱ > q_۱ = x - \frac{< x, ۱ >}{< ۱, ۱ >} ۱ = x - \frac{\int_۰^۱ x dx}{\int_۰^۱ ۱ dx} = x - \frac{\frac{x^۲}{۲}|_۰^۱}{x|_۰^۱} = x - \frac{۱}{۲} .$$

$$< x - \frac{۱}{۲}, x - \frac{۱}{۲} >= \int_۰^۱ (x^۲ - x + \frac{۱}{۴}) dx = \frac{x^۳}{۳} - \frac{x^۲}{۲} + \frac{۱}{۴} x|_۰^۱ = \frac{۱}{۳} - \frac{۱}{۲} + \frac{۱}{۴} = \frac{۱}{۱۲} .$$

پس $< ۱, \sqrt{۱۲}(x - \frac{۱}{۲}) >$ پایه‌ای متعامد یکه برای W است و در نتیجه، بهترین تقریب خطی $x^۵$ برابر است با:

$$C + Dx = \frac{< x^۵, ۱ >}{< ۱, ۱ >} ۱ + \frac{< x^۵, x - \frac{۱}{۲} >}{< x - \frac{۱}{۲}, x - \frac{۱}{۲} >}(x - \frac{۱}{۲}) = \frac{۱}{۶} + \frac{۵}{۷}(x - \frac{۱}{۲})$$

یادداشت: توجه کنید مثال قبل را به گونه‌ای دیگر نیز می‌توان محاسبه نمود که معادل تعریف یافتن بهترین تقریب است. برای یافتن $C + Dx$ بایستی D و C را به گونه‌ای بیابیم که فاصله‌ی بین $x^۵$ و $C + Dx$ کمینه شود؛ یعنی $\| x^۵ - C - Dx \|^۲$ کمینه شود. بنابراین اگر قرار دهیم $F(C, D) = \| x^۵ - C - Dx \|^۲$، باید C و D را به گونه‌ای بیابیم که $F(C, D)$ کمینه شود:

$$F(C, D) = \int_۰^۱ (x^۵ - C - Dx)^۲ dx = \frac{۱}{۱۱} - \frac{۲}{۶} C - \frac{۲}{۷} D + C^۲ + CD + \frac{۱}{۳} D^۲$$

$$\frac{\partial F}{\partial C} = -\frac{۲}{۶} + ۲C + D = ۰ \quad \Rightarrow \quad C + \frac{۱}{۲} D = \frac{۱}{۶}$$

$$\frac{\partial F}{\partial D} = -\frac{۲}{۷} + C + \frac{۲}{۳} D = ۰ \quad \Rightarrow \quad \frac{۱}{۲} C + \frac{۱}{۳} D = \frac{۱}{۷}$$

$$\Rightarrow \begin{bmatrix} ۱ & \frac{۱}{۲} \\ \frac{۱}{۲} & \frac{۱}{۳} \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} \frac{۱}{۶} \\ \frac{۱}{۷} \end{bmatrix} .$$

۷
دترمینان

فرض کنید متوازی‌السطوح n بعدی که توسط بردارهای $a_۱$ تا a_n در V ساخته شده، داده شده است. می‌خواهیم عدد جبری حقیقی‌ای به عنوان حجم متوازی‌السطوح به آن متناظر کنیم؛ به عبارت دقیق‌تر، می‌خواهیم تابع $\Phi : V \times \cdots \times V \to R$ (به تعداد n تا V) را به عنوان نگاشت حجم، معرفی کنیم. انتظاراتی از این تابع داریم که به شرح زیر است:

۱ ـ اگر در راستای یکی از بردارهای a_i، متوازی‌السطوح را به اندازه‌ی λ منبسط کنیم، حجم λ برابر شود.

۲ ـ اگر به یکی از بردارهای a_i (ساق متوازی‌السطوح)، برداری مانند w اضافه شود، حجم متوازی‌السطوح جدید برابر با حجم متوازی‌السطوح اولیه به‌علاوه‌ی حجم متوازی‌السطوحی که ساق‌های آن $a_۱, \cdots, a_{i-۱}, w, a_{i+۱}, \cdots, a_n$ است، شود؛ به عبارت دیگر:

$$\Phi(a_۱, \cdots, a_i+w, a_{i+۱}, \cdots, a_n) = \Phi(a_۱, \cdots, a_i, a_{i+۱}, \cdots, a_n)+\Phi(a_۱, \cdots, w, a_{i+۱}, \cdots, a_n)$$

۳ ـ قدرمطلق حجم متوازی‌السطوح با جابه‌جایی دو ساق ثابت می‌ماند و تنها علامت جبری آن تغییر می‌کند.

۴ ـ حجم متوازی‌السطوح n بعدی روی پایه‌ی استاندارد، برابر یک است.

تعبیر شرایط فوق

۱ ـ تعبیر شرایط اول و دوم: نگاشت حجم Φ، تابعی nـ خطی باشد، یعنی نسبت به هر مولفه‌ی a_i، خطی باشد.

۲ ـ تعبیر شرط سوم: از این شرط به عنوان alternating بودن یاد می‌شود و:

$$\Phi(a_1, \cdots, a_{i-1}, a_i, a_{i+1}, \cdots, a_{j-1}, a_j, a_{j+1}, \cdots, a_n) =$$
$$\Phi(a_1, \cdots, a_{i-1}, a_j, a_{i+1}, \cdots, a_{j-1}, a_i, a_{j+1}, \cdots, a_n) \quad (*)$$

۳ ـ تعبیر شرط چهارم:

$$\Phi(e_1, \cdots, e_n) = 1 .$$

بنابراین می‌خواهیم تابعی nـخطی و alternating یعنی $\Phi : V \times \cdots \times V \to R$ با شرط $\Phi(e_1, \cdots, e_n) = 1$ را به عنوان نگاشت حجم معرفی کنیم.

یادداشت ۱: تابع $\Phi : V \times \cdots \times V \to R$ یک تابع alternating است، اگر و تنها اگر

$$\Phi(a_1, \cdots, a_{i-1}, v, a_{i+1}, \cdots, a_{j-1}, v, a_{j+1}, \cdots, a_n) = 0$$

یادداشت ۲: مجموعه‌ی همه‌ی جایگشت‌های اعداد ۱ تا n را با S نمایش می‌دهیم.

یادداشت ۳: توجه کنید که اگر $a_i = \sum_{j_1=1}^{n} a_{ij_1} e_{j_1}$ آن‌گاه n خطی بودن Φ ایجاب می‌کند که

$$\Phi(a_1, \cdots, a_n) = \Phi(\sum_{j_1=1}^{n} a_{1j_1} e_{j_1}, \cdots, \sum_{j_n=1}^{n} a_{nj_n} e_{j_n}) = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} \cdots a_{nj_n} \Phi(e_{j_1}, \cdots, e_{j_n})$$

از طرفی alternating بودن ایجاب می‌کند که $\Phi(e_{j_1}, \cdots, e_{j_n}) = 0$ اگر و تنها اگر حداقل وجود داشته باشد j_k و j_s ای که $j_k = j_s$. بنابراین، تنها $\Phi(e_{j_1}, \cdots, e_{j_n})$ هایی احتمالا ناصفرند که $j_1, \cdots, j_n = 1, \cdots, n$. در نتیجه، تنها $\Phi(e_{j_1}, \cdots, e_{j_n})$ هایی در معادله‌ی (**) دارای اهمیت‌اند که اندیس e_{j_k} ها به ازای $1 \leq j_k \leq n$، یک جایگشت روی مجموعه‌ی اعداد ۱ تا n بدهند. توجه کنید که $\sigma : \{1, \cdots, n\}^n \to \{1, \cdots, n\}$ یک جایگشت است، اگر و تنها اگر $\{\sigma(1), \cdots, \sigma(n)\} = \{1, \cdots, n\}$؛ بنابراین:

$$\Phi(a_1, \cdots, a_n) = \sum \sigma \in S_n a_{1\sigma(1)} \cdots a_{n\sigma(n)} \Phi(e_{\sigma(1)}, \cdots, e_{\sigma(n)})$$

بار دیگر، می‌توان از خاصیت alternating بودن استفاده کرد و $\Phi(e_{\sigma(1)}, \cdots, e_{\sigma(n)})$ را تبدیل به $\Phi(e_1, \cdots, e_n)$ کرد؛ برای این کار، باید اندیس e_1 در $\Phi(e_{\sigma(1)}, \cdots, e_{\sigma(n)})$ را با $e_{\sigma(1)}$

جابه‌جا کنیم؛ بنابراین بایستی جایگشت دوتایی (i, j) (که فرض کنید j مکان e_i در $\Phi(e_{\sigma(۱)}, \cdots, e_{\sigma(n)})$ است) را مطالعه کنیم. این جایگشت‌ها که به غیر از دو عنصر از مجموعه اعداد ۱ تا n، بقبه ثابت می‌مانند را **ترانهش**[۱] می‌نامند.

تعریف: فرض کنید $X = \{۱, \cdots, n\}$. مجموعه‌ی همه‌ی توابع یک‌به‌یک و پوشا از X به X را مجموعه‌ی جایگشت‌های روی X می‌نامیم و با نماد S_n نمایش می‌دهیم.
واضح است که $|S_n| = n!$ (که در آن، $|S_n|$ تعداد اعضای مجموعه‌ی S_n است). فرض کنید $\sigma \in S_n$ یعنی $\sigma : X \to X$ یک تابع یک به یک و پوشا باشد. معمولاً $\sigma \in S_n$ را به صورت

$$\sigma = \begin{pmatrix} ۱ & ۲ & ۳ & \cdots & n \\ \sigma(۱) & \sigma(۲) & \sigma(۳) & \cdots & \sigma(n) \end{pmatrix}$$

نمایش می‌دهند. به طور مثال:

$$\sigma = \begin{pmatrix} ۱ & ۲ & ۳ \\ ۲ & ۳ & ۱ \end{pmatrix}$$

به این معنی است که $\sigma : \{۱, ۲, ۳\} \to \{۱, ۲, ۳\}$ به طوری که $\sigma(۱) = ۲$، $\sigma(۲) = ۳$، $\sigma(۳) = ۱$.

تعریف: فرض کنید $\sigma \in S_n$ به طوری که وجود داشته باشد $x_۱, \cdots, x_r \in \{۱, \cdots, n\}$ به طوری که:

$$\delta(i) = \begin{cases} i + ۱ & i \in \{x_۱, \cdots, x_{r-۱}\} \\ ۱ & i = x_r \\ i & i \notin \{x_۱, \cdots, x_r\} \end{cases}$$

در این صورت σ **دور** به طول r نامیده شده و با نماد $(x_۱, \cdots, x_r)$ نمایش داده می‌شود.
در نتیجه، تعداد کل دورها برابر است با:

$$\binom{n}{r} \times (r-۱)! = \frac{n!}{r!(n-r)!}(r-۱)! = \frac{n!}{r(n-r)!}$$

تعریف: هر دور به طول دو را یک **ترانهش**[۲] گویند.

مثال: $\sigma = (۱, ۲)$ یک ترانهش در S_n است که $\sigma(۱) = ۲$ و $\sigma(۲) = ۱$ و به ازای هر $۳ \leq i \leq n$ $\sigma(i) = i$.

[۱] transposition [۲] Transposition

مثال: مجموعه‌ی $S_۳$ را در نظر بگیرید؛ $|S_۳| = ۶$ که اعضای آن، به صورت زیر هستند:

$$\begin{bmatrix} ۱ & ۲ & ۳ \\ ۱ & ۲ & ۳ \end{bmatrix}, \quad \begin{bmatrix} ۱ & ۲ & ۳ \\ ۳ & ۱ & ۲ \end{bmatrix}, \quad \begin{bmatrix} ۱ & ۲ & ۳ \\ ۱ & ۳ & ۲ \end{bmatrix}$$

$$\begin{bmatrix} ۱ & ۲ & ۳ \\ ۳ & ۲ & ۱ \end{bmatrix}, \quad \begin{bmatrix} ۱ & ۲ & ۳ \\ ۲ & ۱ & ۳ \end{bmatrix}, \quad \begin{bmatrix} ۱ & ۲ & ۳ \end{bmatrix}$$

$$\begin{bmatrix} ۱ & ۳ & ۲ \end{bmatrix}, \quad \begin{bmatrix} ۲ & ۳ \end{bmatrix}, \quad \begin{bmatrix} ۱ & ۳ \end{bmatrix}, \quad \begin{bmatrix} ۱ & ۲ \end{bmatrix}, \quad \begin{bmatrix} ۱ & ۳ \end{bmatrix}\begin{bmatrix} ۱ & ۲ \end{bmatrix}$$

و همانی.

قضیه: هر جایگشت $\sigma \in S_n$ را می‌توان به حاصل‌ضربی از دورهای دوبه‌دو مجزا تجزیه کرد و صرف نظر از ترتیب دورهای این تجزیه، تجزیه‌ی حاصل‌ضرب دورهای مجزا یکتاست.

طرح اثبات: $i \in \{۱, \cdots, n\}$ را در نظر بگیرید. فرض کنید r کوچکترین عدد طبیعی باشد که $\sigma^r(i) = i$.

زیرا $\{i, \sigma(i), \sigma^۲(i), \cdots, \sigma^n(i)\} \subseteq \{۱, \cdots, n\}$. بنابراین بنا به اصل لانه‌ی کبوتری وجود دارد $۰ \leq r, s \leq n$ که $t \neq s$ به طوری که $\sigma^t(i) = \sigma^s(i)$ (توجه کنید که $i = \sigma^۰(i)$ فرض شده است). بنابراین، با فرض این‌که $t < s$، خواهیم داشت $\sigma^{st}(i) = i$.

حال دور rـ تایی $(i, \sigma(i), \cdots, \sigma^{r-۱}(i))$ را در نظر بگیرید.

- تحدید σ روی مجموعه‌ی $\{۱, \cdots, n\} \setminus \{i, \sigma(i), \cdots, \sigma^{r-۱}(i)\}$ یک جایگشت روی A است.

- با استقرا روی تعداد اعضای مجموعه، جایگشت روی A را می‌توان به حاصل‌ضرب دورهای دوبه‌دو مجزا تجزیه کرد.

- لذا می‌توان σ را به دورهای دوبه‌دو مجزا تجزیه کرد.

در مرحله‌ی بعدی اثبات یکتایی، فرض می‌کنیم که $\sigma = \sigma_۱ \cdots \sigma_r$ و $\sigma = z_۱ \cdots z_s$ نوشته شده است. ادعا می‌کنیم $r = s$ و به ازای هر σ_i یک z_j وجود دارد که $\sigma_i = z_j$. برای اثبات از این‌که $z_۱ = (i, \sigma(i), \cdots, \sigma^{t-۱}(i))$ است استفاده می‌شود.

نتیجه: هر جایگشتی را می‌توان به حاصل‌ضربی از ترانهش‌ها تجزیه کرد.

برهان: کافی است نشان دهیم هر دوری را می‌توان به صورت حاصل‌ضرب ترانهش‌ها نوشت. فرض کنید $(a_1, \cdots, a_m)$ یک دور باشد. در این صورت:

$$(a_1 \cdots a_m) = (a_1 \quad a_2)(a_2 \quad a_3)\cdots(a_{m-1} \quad a_m)$$

زیرا قرار دهید $\sigma = (a_1 \cdots a_m)$ و $z_i = (a_i \quad a_{i+1})$ به ازای هر $1 \le i \le m-1$. فرض کنید $1 \le j \le m$ و $\sigma(a_j)$ و $z_1 \cdots z_m(a_j)$ را به ازای هر j محاسبه می‌کنیم:
اگر $1 \le j < m$ آنگاه $\sigma(a_j) = a_{j+1}$

$$z_1 \cdots z_{j-1} z_j z_{j+1} \cdots z_{m-1}(a_j) = z_1 \cdots z_{j-1} z_j(a_j) = z_1 \cdots z_{j-1}(a_{j+1}) = a_{j+1}$$

یادداشت ۱: تجزیه به ترانهش‌ها یکتا نیست؛ علاوه بر تجزیه‌ای که پیش‌تر معرفی شد، تجزیه‌ی زیر هم برقرار است:

$$(a_1 \cdots a_m) = (a_1 \quad a_m)(a_1 \quad a_{m-1})\cdots(a_1 \quad a_2)$$

درستی‌آزمایی: قرار دهید $\sigma = (a_1 \cdots a_m)$ و $z_i = (a_1 \quad a_i)$ به ازای $2 \le i \le m$. فرض کنید $1 \le j \le m$ و $\sigma(a_j) = a_{j+1}$؛ در نتیجه، خواهیم داشت:

$$z_m \cdots z_{j+1} z_j z_{j-1} \cdots z_2(a_j) = z_m \cdots z_j(a_j) = z_m \cdots z_{j+1}(a_1) = z_m \cdots z_{j+2}(a_{j+1}) = a_{j+1}$$

یادداشت ۲: یک جایگشت در S_n را لزوماً نمی‌توان به صورت تعدادی ترانهش مجزا تجزیه کرد. جایگشت $(1 \quad 2 \quad 3)$ را در S_3 را در نظر بگیرید:

- $(1 \quad 2 \quad 3) \neq (a \quad b)$ زیرا جایگشت سمت چپ سه نقطه را حرکت می‌دهد، در صورتی که سمت راست دو نقطه را.

- $(1 \quad 2 \quad 3) \neq (a \neq b)(c \quad d)$ زیرا به طور مشابه سمت چپ سه نقطه و سمت راست ۴ نقطه را تغییر می‌دهد.

قضیه: اگر $\sigma \in S_n$ را بتوان هم به حاصل‌ضرب r ترانهش و هم حاصل‌ضرب s ترانهش نوشت، آنگاه r و s یا هر دو زوج هستند و یا هر دو فرد.
تعریف: جایگشت $\sigma \in S_n$ را زوج گوییم هر گاه بتوان آن را به صورت حاصل‌ضرب تعداد زوجی

ترانهش نوشت و σ را فرد نامیم هر گاه بتوان آن را به صورت تعداد فردی ترانهش نوشت.
یادداشت ۳: تابع علامت، sgn، روی S_n را به صورت زیر تعریف می‌کنیم:

$$sgn : S_n \to \{\mp 1\}$$

$$sgn(\sigma) = \begin{cases} +1 & \text{اگر } \sigma \text{ زوج باشد} \\ -1 & \text{اگر } \sigma \text{ فرد باشد} \end{cases}$$

یادداشت ۴: هر دور فرد، یک جایگشت زوج و هر دور زوج، یک جایگشت فرد است، زیرا:

$$(a_1 \cdots a_m) = (a_1 \quad a_2)(a_2 \quad a_3) \cdots (a_{m-1} \cdots a_m)$$

رده‌بندی همه‌ی توابع nـ خطی و alternating روی فضای خطی V

با

$$\Phi : V \times \cdots \times V \to R$$

به‌طوری‌که

$$\Phi(a_1, \cdots, a_i, \cdots, a_j, \cdots, a_n) = -\Phi(a_1, \cdots, a_j, \cdots, a_i, \cdots, a_n)$$

با استفاده از خاصیت nـ خطی و alternating بودن Φ نشان دادیم که

$$\Phi(a_1, \cdots, a_n) = \sum_{\sigma \in s_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \Phi(e_{\sigma(1)}, \cdots, e_{\sigma(n)})$$

بنا به قضیه، $\sigma \in s_n$ را می‌توان به صورت دورهای مجزا تجزیه کرد. فرض کنید $\sigma = \sigma_1 \cdots \sigma_s$ (که **مجزا** هستند) و σ_i به ازای $1 \leq i \leq s$ را نیز می‌توان به حاصل‌ضرب ترانهش‌ها تجزیه کرد. بنابراین اگر $\sigma(j) = 1$، باید اولا i, j دقیقا با هم در یک دور σ_t به ازای $1 \leq t \leq s$ ظاهر شود (چون دورهای مجزا هستند) و ثانیا، دور $\sigma_t = (a_1, \cdots, a_m)$ باید به تعداد ترانهش‌هایی که در تجزیه‌ی σ_t ظاهر می‌شود، جابه‌جا شوند (به ازای هر $1 \leq t \leq s$)؛ بنابراین

$$\Phi(e_{\sigma(1)} \cdots e_{\sigma(n)}) = sgn(\sigma)\Phi(e_1, \cdots, e_n) \, .$$

مثال:

$$\Phi(e_۴\ e_۲\ e_۱\ e_۳)$$

$$\sigma = \begin{bmatrix} ۱ & ۲ & ۳ & ۴ \\ ۴ & ۲ & ۱ & ۳ \end{bmatrix} = (۲)(۱\ ۴\ ۳) = ۲(۱\ ۴)(۴\ ۳)$$

در حاشیه، توجه کنید که برای هر n میان ۱ و ۴ و نیز نامساوی ۲،

$$(۲) = (۲\ n)(n\ ۲)\ .$$

بنابراین

$$\Phi(a_۱, \cdots, a_n) = \sum_{\sigma \in S_n} a_{۱\sigma(۱)} \cdots a_{n\sigma(n)} sgn(\sigma) \Phi(e_۱ \cdots e_n)\ .$$

قضیه (رده‌بندی)

اگر V فضای خطی و $B = \{e_۱, \cdots, e_n\}$ پایه‌ای برای V باشد، هر تابع nـ خطی

$$\Phi : V \times \cdots \times V \to R$$

که alternating است، با مقدار $\Phi(e_۱, \cdots, e_n)$ به طور یکتا مشخص می‌شود؛ به عبارت دقیق‌تر،

$$\Phi(a_۱, \cdots, a_n) = \sum_{\sigma \in S_n} a_{۱\sigma(۱)} \cdots a_{n\sigma(n)} sgn(\sigma) \Phi(e_۱ \cdots e_n)$$

که در آن به ازای هر i از ۱ تا n داریم

$$a_i = \sum_{j_i = ۱}^{n} a_{ij_i} e_{j_i}\ .$$

یادداشت: در تعریف نگاشت حجم، شرط $\Phi(e_۱ \cdots e_n) = ۱$ نیز در نظر گرفته شد؛ بنابراین

$$\Phi(a_۱, \cdots, a_n) = \sum_{\sigma \in S_n} sgn(\sigma) a_{۱\sigma(۱)} \cdots a_{n\sigma(n)}\ .$$

یادداشت: مختصات بردار a_i به ازای i از ۱ تا n را به عنوان سطر i ام ماتریس A در نظر بگیرید، به عبارت دیگر،

$$A = \begin{bmatrix} A_۱ \\ \vdots \\ A_n \end{bmatrix}$$

و

$$A_i = ([a_i]_B)^T$$

در این صورت، دترمینان A را برابر با مقدار حجم متوازی‌السطوح روی $a_۱, \cdots, a_n$ تعریف می‌کنیم:

$$detA = \sum_{\sigma \in S_n} sgn(\sigma) a_{۱\sigma(۱)} \cdots a_{n\sigma(n)}$$

یادداشت: چون اعضای S_n همچون σ، جایگشت روی مجموعه‌ی اعداد ۱ تا n هستند، بنابراین $\{a_{۱\sigma(۱)} \cdots a_{n\sigma(n)}\}$ در اصل n تا درایه از $n^۲$ درایه‌ی A هستند که هیچ دو تایی در یک سطر و یک ستون، مشترک نیستند. این مجموعه‌ی n تایی از درایه‌ها، قطر پراکنده نام دارد. به وضوح هر ماتریس $A \in M_n(R)$، دارای $n!$ قطر پراکنده است که این تعداد برابر تعداد جایگشت‌ها روی مجموعه‌ی اعداد ۱ تا n نیز است. به عبارت دیگر، به هر قطر پراکنده‌ی $A \in M_n(R)$، می‌توان یک جایگشت در S_n نسبت داد و برعکس:

$$A = \begin{bmatrix} ۱ & ۵ & ۷ \\ ۸ & ۰ & ۲ \\ ۳ & ۱ & ۴ \end{bmatrix}$$

دترمینان خاصیت‌های زیر را داراست:

- جابه‌جایی دو سطر A علامت دترمینان را تغییر می‌دهد.
- $det\ I = ۱$ که در آن I ماتریس همانی است.
- دترمینان تابعی n ـ خطی نسبت به هر سطر است.

- اگر ماتریس A دو سطر یکسان داشته باشد، آنگاه $det\ A = ۰$.

- کم کردن ضریب یک سطر از سطر دیگر، مقدار دترمینان را تغییر نمی‌دهد. فرض کنید $A = \begin{bmatrix} A_۱ \\ \vdots \\ A_n \end{bmatrix}$. بنابراین:

$$det\ A = \Phi(A_۱, \cdots, A_n)$$

$$\Phi(A_۱, \cdots, A_i - cA_j, A_{i+۱}, \cdots, A_j, \cdots, A_n)$$

$$= \Phi(A_۱, \cdots, A_i, A_{i+۱}, \cdots, A_j, \cdots, A_n) - c\,\underbrace{\Phi(A_۱, \cdots, A_j, A_{i+۱}, \cdots, A_j, \cdots, A_n}_{۰}$$

$$= \Phi(A_۱, \cdots, A_n) = det\ A$$

- اگر یکی از سطرهای A صفر باشد، $det\ A = ۰$. زیرا:

$$det\ A = \sum_{\sigma \in S_n} sgn(\sigma) a_{۱\sigma(۱)} \cdots a_{n\sigma(n)} = ۰$$

بنابراین همواره در قطر پراکنده‌ی معادل σ یک صفر ظاهر می‌شود.

- اگر A ماتریسی مثلثی باشد، آنگاه:

$$det\ A = a_{۱۱} \cdots a_{nn}$$

زیرا در هر جایگشتی به غیر از جایگشت همانی، عنصری از زیر قطر اصلی ظاهر می‌شود که اگر A بالامثلثی باشد، آنگاه جمله‌ی متناظر با σ – یعنی $(a_{۱\sigma(۱)} \cdots a_{n\sigma(n)})$ برابر صفر خواهد بود.

- اگر A تکین باشد، آنگاه $det\ A = ۰$. فرض کنید A تکین است. پس سطرهای A وابسته‌ی خطی هستند. یعنی اگر $A = \begin{bmatrix} A_۱ \\ \vdots \\ A_n \end{bmatrix}$، آنگاه ترکیب خطی از $A_۱, \cdots, A_n$ مانند $c_۱A_۱ + \cdots + c_nA_n = ۰$ وجود دارد به طوری که حداقل یکی از c_i ها ناصفر است.

فرض کنید $c_i \neq ۰$. بنابراین:

$$A_i = -(\frac{c_۱}{c_i}A_۱ + \cdots + \frac{c_{i-۱}}{c_i}A_{i-۱} + \frac{c_{i+۱}}{c_i}A_{i+۱} + \cdots + \frac{c_n}{c_i}A_n)$$

از آنجایی که det تابعی n ـ خطی نسبت به سطرها است، پس:

$$det\ A = \sum_{j=۱, j\neq i}^{n} \frac{-c_j}{c_i} det \begin{bmatrix} A_۱ \\ \vdots \\ A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} = ۰$$

بنابراین اگر A تکین باشد، آنگاه $det\ A = ۰$. پس اگر $det\ A \neq ۰$ آنگاه A وارون‌پذیر است.

- $det\ A = det\ A^T$

هر جمله‌ی $sgn(\sigma)a_{۱\sigma(۱)} \cdots a_{n\sigma(n)}$ در $detA$ معادل $sgn(\sigma^{-۱})a_{۱\sigma^{-۱}(۱)} \cdots a_{n\sigma^{-۱}(n)}$ در $det\ A^T$ است. به عبارت دیگر اگر $\sigma(i) = j$، آنگاه $i = \sigma^{-۱}(j)$:

$$det\ A^T = \sum_{\sigma \in S_n} sgn(\sigma^{-۱})a_{۱\sigma^{-۱}(۱)} \cdots a_{n\sigma^{-۱}(n)}$$

جای سطر و ستون عوض شد. اگر ثابت کنیم $sgn(\sigma) = sgn(\sigma^{-۱})$، اثبات کامل می‌شود. فرض کنید $\sigma = \tau_۱ \cdots \tau_m$ که در آن τ_i ترانهش است. بنابراین $\sigma^{-۱} = \tau_m^{-۱} \cdots \tau_۱^{-۱}$ و $sgn(\sigma) = sgn(\sigma^{-۱})$.

- فرض کنید $A, B \in M_n(R)$. در این صورت:

$$det\ AB = (det\ A)(det\ B)$$

برهان: با فرض اینکه سطرهای A با $a_۱, \cdots, a_n$ نمایش داده شده، تابع Φ را تعریف

می‌کنیم:

$$\Phi : V \times \cdots \times V \to R$$

$$\Phi(a_۱ \cdots a_n) = det\ AB$$

ابتدا، نشان می‌دهیم تابع Φ n ـ خطی است.

به ازای ماتریس $A = \begin{bmatrix} a_۱ \\ \vdots \\ a_n \end{bmatrix}$:

$$det\ AB = det\left(\begin{bmatrix} a_۱ \\ \vdots \\ a_i \\ a_{i+۱} \\ \vdots \\ a_n \end{bmatrix} B\right) = det\begin{bmatrix} a_۱ B \\ \vdots \\ a_i B \\ a_{i+۱} B \\ \vdots \\ a_n B \end{bmatrix}$$

بنابراین چون det تابع n ـ خطی است،

$$det\begin{bmatrix} a_۱ B \\ \vdots \\ (ca_i + a'_i)B \\ a_{i+۱} B \\ \vdots \\ a_n B \end{bmatrix} = cdet\begin{bmatrix} a_۱ B \\ \vdots \\ a_i B \\ a_{i+۱} B \\ \vdots \\ a_n B \end{bmatrix} + det\begin{bmatrix} a_۱ B \\ \vdots \\ a'_i B \\ a_{i+۱} B \\ \vdots \\ a_n B \end{bmatrix}$$

پس تابع Φ n ـ خطی است.

هم‌چنین، تابع Φ متناوب $(alternative)$ است، زیرا فرض کنید $a_i = a_j = a$.

$$\Phi(a_۱, \cdots, a_i, \cdots, a_j, \cdots a_n) = det \begin{bmatrix} a_۱B \\ \vdots \\ aB \\ \vdots \\ aB \\ \vdots \\ a_nB \end{bmatrix} = ۰$$

چرا که det یک تابع متناوب است. بنا به قضیه‌ی رده‌بندی توابع n ـ خطی و $alternating$:

$$\Phi(a_۱, \cdots, a_n) = \sum_{\sigma \in S_n} sgn(\sigma) a_{۱\sigma(۱)} \cdots a_{n\sigma(n)} \Phi(e_۱, \cdots, e_n) = (detA)\Phi(e_۱, \cdots, e_n)$$

از طرفی:

$$\Phi(e_۱, \cdots, e_n) = det\ B$$

پس:

$$\Phi(a_۱, \cdots, a_n) = (det\ A)(det\ B) \quad \Rightarrow \quad det\ AB = (det\ A)(det\ B)$$

- اگر A وارون‌پذیر باشد، آنگاه $det\ A \neq ۰$.

$$A \text{وارون‌پذیر} \quad \Rightarrow \quad AA^{-۱} = I \quad \Rightarrow \quad (detA)(detA^{-۱}) = ۱ \quad \Rightarrow \quad detA \neq ۰$$

خاصیت ۱۲: فرض کنید $A \in M_r(R)$ و $C \in M_s(R)$ و $B \in M_{rs}(R)$؛ در این صورت

$$det \begin{bmatrix} A & B \\ ۰ & C \end{bmatrix} = (detA)(detC)$$

برهان: ماتریس‌های A و B را دو ماتریس ثابت فرض کنید و قرار دهید $V = R^s$ و نگاشت زیر را تعریف کنید:

$$\Phi : V \times \cdots \times V \to R$$
$$\to \Phi(c_1, \cdots, c_s) = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

که در آن، سطر i ام C برابر c_iاست. Φ نگاشتی sـ خطی و alternating است (چرا؟)، بنابراین بنا به رده‌بندی این دسته از نگاشت‌های خطی،

$$\Phi(c_1, \cdots, c_s) = detC\Phi(e_1, \cdots, e_s)$$
$$\Phi(e_1, \cdots, e_s) = det \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} = det \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$$

زیرا عملیات سطری، دترمینان را تغییر نمی‌دهد.

با استفاده از تعریف دترمینان (در نظر گرفتن قطر پراکنده)، به‌دست می‌آوریم که

$$det \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} = detA$$

بنابراین

$$\Phi(c_1, \cdots, c_s) = det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = (detC)(detA)$$

خاصیت ۱۳: فرض کنید $A, B, C, D \in M_n(R)$ به طوری‌که $CD = DC$؛ در این صورت،

$$det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = det(AD - BC)$$

برهان:

- حالت اول ــ فرض کنید که D وارون‌پذیر است:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix} = \begin{bmatrix} A - BD^{-1} & B \\ 0 & D \end{bmatrix}$$

در نتیجه

$$det\begin{bmatrix}A & B\\ C & D\end{bmatrix} det\begin{bmatrix}I & ۰\\ -D^{-۱}C & I\end{bmatrix} = det\begin{bmatrix}A-BD^{-۱}C & B\\ ۰ & D\end{bmatrix}$$

$$\rightarrow det\begin{bmatrix}A & B\\ C & D\end{bmatrix}(detI)^{۲} = (det(A-BD^{-۱}C))(detD)$$

بنابراین،

$$det\begin{bmatrix}A & B\\ C & D\end{bmatrix} = det(AD-BD^{-۱}CD),\ CD=DC$$

بنابراین اگر D وارون‌پذیر باشد،

$$det\begin{bmatrix}A & B\\ C & D\end{bmatrix} = det(AD-BC)$$

- حالت دوم ـ فرض کنید که D دل‌خواه است. ماتریس $D+xI$ را در نظر بگیرید.

$$det(D+xI) = det\begin{bmatrix}d_{۱۱}+x & d_{۱۲} & \cdots & d_{۱n}\\ d_{۲۱} & d_{۲۲}+x & \cdots & d_{۲n}\\ \vdots & \vdots & \vdots & \vdots\\ d_{n۱} & d_{n۲} & \cdots & d_{nn}+x\end{bmatrix}$$

با استفاده از تعریف دترمینان، $det(D+xI)$ یک چندجمله‌ای از درجه‌ی n بر حسب x است که ضریب x^nدر آن، یک است، زیرا تنها قطر پراکنده‌ای که x^n را می‌سازد، قطر اصلی است:

$$det(D+xI) = x^n + \alpha_{n-۱}x^{n-۱} + \cdots + \alpha_۱ x + \alpha_۰$$

که در آن a_i ها، اعداد حقیقی هستند.

می‌دانیم که هر چندجمله‌ای از درجه‌ی n، حداکثر n ریشه دارد. فرض کنید S مجموعه‌ی

ریشه‌های آن باشد؛ در این صورت، به ازای هر $x \in S$ (یعنی هر x که ریشه‌ی چندجمله‌ای فوق باشد)، $det(D+xI) \neq ۰$، یعنی به ازای هر $x \in R\setminus S$ ماتریس $D+xI$ وارون‌پذیر است؛ از طرفی، $(D+xI)C = C(D+xI)$ و در نتیجه،

$$det \begin{bmatrix} A & B \\ C & D+xI \end{bmatrix} = det(A(D+xI) - BC)$$

قرار دهید:

$$f(x) = det \begin{bmatrix} A & B \\ C & D+xI \end{bmatrix} = det(A(D+xI) - BC)$$

آن‌گاه، $f(x)$ چندجمله‌ای درجه‌ی n است که به ازای هر $x \in R \setminus S$ داریم $f(x) = ۰$؛ پس چندجمله‌ای f باید چندجمله‌ای صفر باشد، زیرا نامتناهی ریشه دارد. از $f(x) = ۰$ به دست می‌آید که به ازای هر $x \in R$،

$$det \begin{bmatrix} A & B \\ C & D+xI \end{bmatrix} = det(A(D+xI) - BC)$$

بنابراین به ازای $x = ۰$،

$$det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = det(AD - BC)$$

بسط دترمینان نسبت به ستون

گزاره: فرض کنید A ماتریسی $n \times n$ است، در این صورت:

$$det\ A = \sum_{i=۱}^{n} (-۱)^{i+j} a_{ij} det\ A(i|j)$$

برهان: تابع $\Phi : R^n \times R^n$ با ضابطه‌ی $\Phi(a_۱, \cdots, a_n) = \sum_{i=۱}^{n} (-۱)^{i+j} a_{ij} det\ A(i|j)$ را

در نظر بگیرید. تابع Φ ، n ـ خطی و متناوب است. لذا بنا به قضیه‌ی رده‌بندی:

$$\Phi(a_1, \cdots, a_n) = (det\ A)\Phi(e_1, \cdots, e_n)$$

از طرفی $\Phi(e_1, \cdots, e_n) = 1$، لذا حکم ثابت می‌شود.

اثبات شهودی:

$$det\ A = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

حال جملات ظاهرشده در در جمع فوق را بر حسب درایه‌های ستون jام مرتب می‌کنیم:

$$det\ A = \beta_{i1}a_{i1} + \cdots + \beta_{ij}a_{ij} + \cdots + \beta_{in}a_{in}$$

و نشان می‌دهیم که $\beta_{ij} = (-1)^{i+j} det\ A(i|j)$. در رابطه‌ی اول در بخش اثبات شهودی $\beta_{ij}a_{ij}$ جمع روی همه‌ی قطر پراکنده‌هایی است که از سطر iام a_{ij} انتخاب شده است. به عبارت دیگر $\sigma_i = j$ به ازای $\sigma \in S_n$. بقیه‌ی عناصر قطر پراکنده نیز از درایه‌های ماتریس $A(i|j)$ انتخاب می‌شود. با توجه به اینکه $det\ A = det\ A^T$ و خاصیت متناوب بودن دترمینان، با جابه‌جایی a_{ij} و انتقال آن به a_{11}، دترمینان در $(-1)^{(i-1)(j-1)}$ ضرب خواهد شد. زیرا $j - 1$ و $i - 1$ جابه‌جایی ستونی و سطری لازم است. پس:

$$\beta_{ij} = (-1)^{i+j} det\ A(i|j)$$

تعریف: قرار دهید $c_{ij} = (-1)^{i+j} det\ A(i|j)$ و c_{ij} را همساز ijام ماتریس A گویند.

یادداشت: بنا به تعریف:

$$det\ A = \sum_{i=1}^{n} a_{ij}c_{ij}$$

به عبارت دیگر اگر قرار دهید $C = \begin{bmatrix} c_{11} & \cdots & c_{ij} & \cdots & c_{1n} \\ \vdots & & \vdots & & \vdots \\ c_{n1} & \cdots & c_{nj} & \cdots & c_{nn} \end{bmatrix}$ در این صورت $det\, A$ ضرب داخلی ستون jام A در ستون jام C است.

یادداشت: ادعا می‌کنیم اگر $k \neq j$ آنگاه $\sum_{i=1}^{n} a_{ik}c_{ij} = 0$.

برای اثبات، ماتریس B را چنان بسازید که B همان ماتریس A است که ستون jام آن، ستون

kام A تکرار شده است. بنابراین $det\ B = ۰$. از طرفی $det\ B = \sum_{i=۱}^{n} a_{ik}c_{ij}$. بنابراین $\sum_{i=۱}^{n} a_{ik}c_{ij} = ۰$ به ازای $j \neq k$. ماتریس B به ازای $k < j$:

$$B = \begin{bmatrix} a_{۱۱} & \cdots & a_{۱k} & \cdots & a_{۱k} & \cdots & a_{۱n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n۱} & \cdots & a_{nk} & \cdots & \underbrace{a_{nk}}_{\text{ستون}\,j\,\text{ام}} & \cdots & a_{nn} \end{bmatrix}$$

در نتیجه:

$$C^T A = \begin{bmatrix} c_{۱۱} & \cdots & c_{n۱} \\ c_{۱۲} & \cdots & c_{n۲} \\ \vdots & & \vdots \\ c_{۱n} & \cdots & c_{nn} \end{bmatrix} \begin{bmatrix} a_{۱۱} & \cdots & a_{n۱} \\ a_{۲۱} & \cdots & a_{۲n} \\ \vdots & & \vdots \\ a_{n۱} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} det\ A & & ۰ \\ & \ddots & \\ ۰ & & det\ A \end{bmatrix} = (detA)I$$

تعریف: ماتریس C^T را ماتریس الحاقی کلاسیک ماتریس A گویند و با نماد $adj\,A$ نمایش می‌دهند. پس $(adj\ A)A = (det\ A)I$.

خواص:

- $(adj\ A)^T = adj\ A^T$. زیرا:

$$(adj A^T)_{ij} = (-۱)^{i+j} detA^T(j|i) = (-۱)^{i+j} detA(i|j) = (adj A)_{ji} = (adj A)^T_{ij}$$

- $(adj\ A)A = A(adj\ A)$. زیرا از رابطه‌ی $(adj\ A)A = (det\ A)I$ داریم:

$$A^T(adj\ A)^T = (det\ A)I$$

پس:

$$A^T adj\ A^T = (det\ A)I$$

از طرفی

$$(adj\ A^T) = (det\ A)I$$

پس:

$$(adj\ A^T)A^T = A^T(adj\ A^T)$$

کافی است قرار دهید $B = A^T$ و به دست می‌آید:

$$(adj\ A^T) = A(adj\ A)$$

- اگر A وارون‌پذیر باشد آنگاه:

$$A(\frac{adj\ A}{det\ A}) = (\frac{adj\ A}{det\ A})A = I \quad \Rightarrow \quad A^{-۱} = (\frac{adj\ A}{det\ A})$$

قاعده‌ی کرامر برای حل دستگاه خطی $Ax = b$

فرض کنید $A \in M_n(R)$ ماتریس وارون‌پذیر باشد. می‌خواهیم دستگاه $Ax = b$ را حل کنیم.

$$Ax = b \quad \Rightarrow \quad (adj\ A)Ax = (adj\ A)b \quad \Rightarrow \quad (det\ A)x = (adj\ A)b$$

$$\Rightarrow \quad (det\ A)x_j = \sum_{i=۱}^{n}(-۱)^{i+j} det\ (A(i|j))b_i$$

بنابراین اگر B_j ماتریسی $n \times n$ باشد که از قرار دادن b به جای ستون jام ماتریس A به دست آمده باشد:

$$B_j = \begin{bmatrix} a_{۱۱} & \cdots & b_۱ & \cdots & a_{۱n} \\ \vdots & & \vdots & & \vdots \\ a_{n۱} & \cdots & b_n & \cdots & a_{nn} \end{bmatrix}$$

در صورت بسط دترمینان B_j نسبت به ستون jام داریم:

$$det\ B_j = \sum_{i=۱}^{n}(-۱)^{i+j} det\ (A(i|j)b_i)$$

در نتیجه به ازای هر j، $۱ \leq j \leq n$، داریم:

$$x_j = \frac{det\ B_j}{det\ A}$$

که به عبارت بالا، قاعده‌ی کرامر گویند.

یادداشت: فرض کنید V یک فضای خطی با بعد متناهی و $T : V \to V$ تبدیل خطی است. مجموعه‌ی B را پایه‌ای برای V در نظر بگیرید. در این صورت دترمینان T را با نماد $det\, T$ نمایش می‌دهیم و تعریف می‌کنیم:

$$det\, T = det\, [T]_B$$

توجه کنید که $det\, T$ مستقل از پایه‌ی B است. زیرا فرض کنید B' پایه‌ی دیگری برای V باشد، در این صورت ماتریس وارون‌پذیر P وجود دارد به طوری که $[T]_B = P[T]_B P^{-۱}$. در نتیجه:

$$det\, [T]_B = (det\, P)(det\, [T]_{B'})(det\, P^{-۱})$$

پس:

$$det\, [T]_B = det\, [T]_{B'}$$

۸

بردارهای ویژه و مقادیر ویژه

مقدمه

فرض کنید دستگاه معادلات دیفرانسیل زیر با مقادیر اولیه‌ی $u_۱(۰) = ۸$ و $u_۲(۰) = ۵$ داده شده است.

$$\frac{du_۱(t)}{dt} = ۴u_۱(t) - ۵u_۲(t)$$
$$\frac{du_۲(t)}{dt} = ۲u_۱(t) - ۳u_۲(t)$$

به یاد دارید که جواب معادله‌ی $\frac{du}{dt} = au(t)$ به ازای $u(۰) = u.$ برابر با $u(t) = e^{at}u(۰)$ است. با پیروی از جواب معادله‌ی تک‌مجهولی، حدس می‌زنیم که $u(t) = ae^{\lambda t}$ که $a = \begin{bmatrix} a_۱ \\ a_۲ \end{bmatrix}$ باید جواب دستگاه فوق باشد، یعنی

$$\lambda e^{\lambda t} a_۱ = ۴e^{\lambda t}a_۱ - ۵e^{\lambda t}a_۲$$
$$\lambda e^{\lambda t} a_۲ = ۲e^{\lambda t}a_۱ - ۳e^{\lambda t}a_۲$$

بنابراین بایستی a و λ را به گونه‌ای بیابیم که

$$\lambda a_۱ = ۴a_۱ - ۵a_۲$$
$$\lambda a_۲ = ۲a_۱ - ۳a_۲$$

و در نتیجه

$$\begin{bmatrix} ۴ & -۵ \\ ۲ & -۳ \end{bmatrix} \begin{bmatrix} a_۱ \\ a_۲ \end{bmatrix} = A \begin{bmatrix} a_۱ \\ a_۲ \end{bmatrix} = \lambda \begin{bmatrix} a_۱ \\ a_۲ \end{bmatrix}$$

به عبارت دیگر، در صورتی‌که بتوان λ و بردار a را به‌گونه‌ای یافت که $Aa = \lambda a$، آن‌گاه $u(t) = ae^{\lambda t}$ جواب دستگاه اولیه خواهد بود. بردار a خاصیت ویژه‌ای دارد: راستای این بردار، تحت A حفظ می‌شود؛ به عبارت دیگر،

$$(A - \lambda I)a = ۰$$

بنابراین، اگر دنبال بردار ناصفری مانند a هستیم که تحت A راستای آن حفظ شود، باید $N(A - \lambda I) \neq \{۰\}$ و به عبارت دیگر، باید λ را به گونه‌ای بیابیم که فضای پوچ $A - \lambda I$ نابدیهی باشد، یعنی باید $A - \lambda I$ تکین باشد، معادلا:

$$det(A - \lambda I) = ۰$$

در مثال فوق،

$$A - \lambda I = \begin{bmatrix} ۴ - \lambda & -۵ \\ ۲ & -۳ - \lambda \end{bmatrix}$$

$$det(A - \lambda I) = (۴ - \lambda)(-۳ - \lambda) + ۱۰ = \lambda^۲ - \lambda - ۲ = ۰$$

$$\rightarrow \lambda_۱ = ۲\,, \quad \lambda_۲ = -۱$$

$$\lambda = -۱ \rightarrow (A - \lambda I)a = \begin{bmatrix} ۵ & -۵ \\ ۲ & -۲ \end{bmatrix} \begin{bmatrix} a_۱ \\ a_۲ \end{bmatrix} = \begin{bmatrix} ۰ \\ ۰ \end{bmatrix} \rightarrow a = \begin{bmatrix} a_۱ \\ a_۲ \end{bmatrix} = \begin{bmatrix} ۱ \\ ۱ \end{bmatrix}$$

$$\rightarrow u(t) = e^{-t} \begin{bmatrix} ۱ \\ ۱ \end{bmatrix} \quad (۲)$$

$$\lambda = ۲ \to (A-\lambda I)a = \begin{bmatrix} ۲ & -۵ \\ ۲ & -۵ \end{bmatrix} \begin{bmatrix} a_۱ \\ a_۲ \end{bmatrix} = \begin{bmatrix} ۰ \\ ۰ \end{bmatrix} \to a = \begin{bmatrix} a_۱ \\ a_۲ \end{bmatrix} = \begin{bmatrix} ۵ \\ ۲ \end{bmatrix}$$

$$\to u(t) = e^{۲t} \begin{bmatrix} ۵ \\ ۲ \end{bmatrix} \quad (۱)$$

جواب‌های (۱) و (۲) جواب‌های خاص معادله‌ی دیفرانسیل هستند، بنابراین مجموعه جواب عمومی معادله‌ی دیفرانسیل فوق، عبارت است از:

$$\{u(t) = c_۱ e^{-t} \begin{bmatrix} ۱ \\ ۱ \end{bmatrix} + c_۲ e^{۲t} \begin{bmatrix} ۵ \\ ۲ \end{bmatrix} \Big| c_۱ \in R\,, \quad c_۲ \in R\}$$

پس جواب معادله‌ی دیفرانسیل، تحت شرایط اولیه‌ی داده شده، عبارت است از:

$$u(t) = ۳e^{-t} \begin{bmatrix} ۱ \\ ۱ \end{bmatrix} + e^{۲t} \begin{bmatrix} ۵ \\ ۲ \end{bmatrix}$$

زیرا، با لحاظ کردن شرایط اولیه‌ی $u_۱(۰) = ۸$ و $u_۲(۰) = ۵$ خواهیم داشت:

$$\begin{bmatrix} ۱ & ۵ \\ ۱ & ۲ \end{bmatrix} \begin{bmatrix} c_۱ \\ c_۲ \end{bmatrix} = \begin{bmatrix} ۸ \\ ۵ \end{bmatrix} \to c_۱ = ۳\,, \quad c_۲ = ۱$$

فرض کنید $T : V \to V$ یک تبدیل خطی روی فضای خطی با بعد متناهی V است. در این صورت به دنبال یافتن بردارهای خاص یا بردارهای ویژه (ناصفر) در فضای خطی V هستیم که تحت تأثیر T راستای آن تغییر نکند، یعنی بردار $x \neq ۰$ که $T(x) = \lambda x$.

گزاره: فرض کنید $T : V \to V$ تبدیل خطی روی فضای متناهی‌البعد V است و λ عدد حقیقی است. در این صورت موارد زیر با هم معادل هستند:

- بردار ناصفر x در فضای خطی V وجود دارد به طوری که $T(x) = \lambda x$.
- تبدیل خطی $T - \lambda I$ وارون‌پذیر نیست.
- $det\, T - \lambda I = ۰$.

برهان:

- $2 \Rightarrow 1$: بردار x در فضای پوچ $T - \lambda I$ است زیرا $(T-\lambda I)x = T(x) - \lambda x = 0$. بنابراین $N(T-\lambda I) \neq 0$. در نتیجه $T - \lambda I$ وارون‌پذیر نیست.

- $3 \Rightarrow 2$: فرض کنید B پایه‌ی فضای خطی V است، اگر $det\ (T-\lambda I) \neq 0$، آنگاه $det\ ([T-\lambda I]_B) \neq 0$. در نتیجه $[T-\lambda I]_B$ ماتریسی وارون‌پذیر است. لذا $T - \lambda I$ تبدیل خطی وارون‌پذیر است که تناقض است.

- $1 \Rightarrow 3$: چون $det\ (T-\lambda I) = 0$، لذا تبدیل خطی $T-\lambda I$ وارون‌پذیر نیست. پس $N(T-\lambda I) \neq \{0\}$. فرض کنید $x \neq 0$ و $x \in N(T-\lambda I)$. پس $(T-\lambda I)x = 0$. یعنی $T(x) = \lambda x$.

تعریف: فرض کنید $A \in M_n(R)$. در این صورت $\lambda \in R$ را یک مقدار ویژه برای A گویند، هر گاه $det\ (\lambda I - A) = 0$. مجموعه‌ی تمام مقادیر ویژه‌ی A را **طیف** A نامند و معمولاً با علامت $spec(A)$ نمایش می‌دهند.

تعریف: فرض کنید $A \in M_n(R)$. چندجمله‌ای ویژه‌ی A را به صورت زیر تعریف می‌کنیم:

$$f(x) = det\ (xI - A)$$

یادداشت ۱: چندجمله‌ای ویژه‌ی A، چندجمله‌ای تکین و از درجه‌ی n است (تکین یعنی ضریب x^n برابر یک است) زیرا:

$$f(x) = det\ (xI - A) = det \begin{bmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ a_{21} & x - a_{22} & & \\ \vdots & & \ddots & \\ -a_{n1} & \cdots & & x - a_{nn} \end{bmatrix}$$

با استفاده از تعریف دترمینان، به راحتی به دست می‌آید که $f(x)$ یک چندجمله‌ای از درجه n است. تنها قطر پراکنده‌ای که جمله‌ی x^n را پدید می‌آورد، قطر اصلی است. پس ضریب x^n برابر با یک است.

یادداشت ۲: ضریب x^{n-1} در چندجمله‌ای ویژه‌ی A را می‌یابیم.

به وضوح جمله‌ی $x^{n-۱}$ فقط در حاصل‌ضرب قطر اصلی که قطر پراکنده است پدید می‌آید. یعنی

$$(x - a_{۱۱}) \cdots (x - a_{nn})$$

بنابراین ضریب $x^{n-۱}$ برابر است با $-a_{۱۱} - a_{۲۲} - \cdots - a_{nn}$. اگر $tr(A)$ را مجموع عناصر روی قطر اصلی A تعریف کنیم، آنگاه ضریب جمله‌ی $x^{n-۱}$ در چندجمله‌ای ویژه‌ی ماتریس A برابر است با $-tr(A)$.

یادداشت ۳: جمله‌ی ثابت چندجمله‌ای ویژه‌ی ماتریس A برابر است با $(-۱)^n det\ A$. زیرا:

$$f(x) = det\ (xI - A) \quad \Rightarrow \quad \text{جمله‌ی ثابت} = f(۰) = det\ (-A) = (-۱)^n det\ A$$

یادداشت ۴: اگر چندجمله‌ای ویژه‌ی ماتریس A به چندجمله‌ای درجه یک تجزیه شود، معادلاً یعنی چندجمله‌ای $f(x)$ دارای n ریشه در اعداد حقیقی باشد، در این صورت دترمینان A برابر با حاصل‌ضرب مقادیر ویژه‌ی A است. زیرا فرض کنید:

$$f(x) = (x - \lambda_۱) \cdots (x - \lambda_n)$$

توجه کنید که ممکن است ریشه‌ی تکراری نیز وجود داشته باشد.

به ازای هر $۱ \leq i \leq n$ داریم $f(\lambda_i) = ۰$. از طرفی، $f(۰) = (-۱)^n \lambda_۱ \cdots \lambda_n$ بنابراین $f(۰) = (-۱)^n detA = (-۱)^n \lambda_۱ \cdots \lambda_n$ و در نتیجه $detA = \lambda_۱ \cdots \lambda_n$. هم‌چنین، ضریب $x^{n-۱}$ نیز برابر با $-(\lambda_۱ \cdots \lambda_n)$ است.

مثال: فرض کنید $A \in M_n(R)$ و قطری است؛ در این صورت، $spec(A)$ را بیابید.

فرض کنید $A = \begin{bmatrix} d_۱ & \cdots & \\ \vdots & \vdots & \\ \cdots & & d_n \end{bmatrix}$ ، در این صورت:

$$f(x) = det(\lambda I - A) = \begin{bmatrix} x - d_۱ & \cdots & \\ \vdots & \vdots & \\ \cdots & & x - d_n \end{bmatrix} = (x - d_۱) \cdots (x - d_n)$$

بنابراین، مجموعه‌ی مقادیر ویژه ($spec(A)$) برابر با $\{d_۱, \cdots, d_n\}$ است و $e_۱$ تا e_n بردارهای

ویژه‌ی متناظر هستند، یعنی به ازای هر $1 \leq i \leq n$ داریم $Ae_i = \lambda_i e_i$.

مثال: مقادیر ویژه‌ی ماتریس افکنش P را بیابید.

فرض کنید $x \neq 0$ و $Px = \lambda x$. می‌دانیم که $P^2 = P$ بنابراین $P^2x = Px = \lambda Px = \lambda^2 x$ پس $\lambda x = \lambda^2 x$ و در نتیجه $(\lambda^2 - \lambda)x = 0$. چون x ناصفر است، پس $\lambda^2 - \lambda = 0$. بنابراین اگر λ مقدار ویژه باشد، آنگاه λ ریشه‌ی چندجمله‌ای $x(x-1)$ است؛ از طرفی، ریشه‌ی چندجمله‌ای $f(x) = det(xI - P)$ مقدار ویژه است، پس

$$f(x) = (\lambda - 1)^r \lambda^{n-r}$$

که در آن، r رتبه‌ی فضای ستونی P است.

مثال: فرض کنید A ماتریسی بالا مثلثی باشد؛ مقادیر ویژه‌ی آن را بیابید.

$$A = \begin{bmatrix} a_{11} & \cdots & \cdots \\ \vdots & \vdots & \vdots \\ 0 & \cdots & a_{nn} \end{bmatrix} \rightarrow det(xI - A) = (x - a_{11}) \cdots (x - a_{nn})$$

بنابراین،

$$spec(A) = \{a_{11}, \cdots, a_{nn}\}$$

مثال: فرض کنید $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in M_2(R)$. مقادیر ویژه‌ی آن را بیابید.

$$f(x) = det(xI - A) = det \begin{bmatrix} x & 1 \\ -1 & x \end{bmatrix} = x^2 + 1$$

مشخصا $f(x)$ ریشه‌ی حقیقی، ندارد، بنابراین A به عنوان ماتریسی در $M_2(R)$، مقدار ویژه ندارد؛ ولی، اگر A را به عنوان ماتریسی در $M_n(C)$ (که در آن C مجموعه‌ی اعداد مختلط است) در نظر بگیریم، مقادیر ویژه دارد، زیرا $x^2 + 1 = (x - i)(x + i)$ و در نتیجه:

- اگر $A \in M_2(R)$ آنگاه $spec(A) = \phi$.
- اگر $A \in M_2(C)$ آنگاه $spec(A) = \{\pm i\}$.

به عبارت دیگر، $M_۲(R)$ فضای ماتریس‌های ۲ در ۲ روی اعداد حقیقی است، یعنی فضای خطی روی اعداد حقیقی به این معنا که اسکالر در آن از اعداد حقیقی انتخاب می‌شود، و $M_۲(C)$ فضای ماتریس‌های ۲ در ۲ با درایه‌های مختلط است و به عنوان فضای خطی، اسکالرهای آن از اعداد مختلط انتخاب می‌شود.

تعریف: فرض کنید $A \in M_n(F)$ که در آن F برابر R یا C است. در این صورت، اگر λ مقدار ویژه‌ی A باشد، آنگاه **فضای ویژه‌ی** مربوط به مقدار ویژه‌ی λ چنین تعریف می‌شود:

$$W = \{x \in F^n | Ax = \lambda x\}$$

یادداشت: توجه شود که W زیرفضای F^n است:

- $۰ \in W \neq \phi$
- اگر $x, y \in W$ و $c \in F$، آنگاه

$$A(cx + y) = cAx + Ay = c\lambda x + \lambda y = \lambda(cx + y)$$

پس $cx + y \in W$.

فضاهای ویژه و ماتریس‌های قطری

تعریف: ماتریس **قطری**، ماتریسی مربعی است که به جز درایه‌های روی قطر، باقی درایه‌ها صفر است.

تعریف: ماتریس $A \in M_n(R)$ را **قطری‌شدنی** گوییم هرگاه ماتریس وارون‌پذیر $S \in M_n(F)$ که $F = R$ یا C وجود داشته باشد به طوری که $S^{-۱}AS$ ماتریس قطری باشد.

مثال: فرض کنید $A = \begin{bmatrix} \frac{۱}{۲} & \frac{۱}{۲} \\ \frac{۱}{۲} & \frac{۱}{۲} \end{bmatrix}$. ماتریسی بیابید که A را قطری کند. به عبارت دیگر S را به گونه‌ای بیابید که $S^{-۱}AS$ قطری باشد.

حل:

- مقادیر ویژه‌ی A را پیدا می‌کنیم:

$$f(\lambda) = det\,(\lambda I - A) = ۰ \quad \Rightarrow \quad \lambda = ۱, ۰$$

- بردار ویژه‌های متناظر با مقادیر ویژه A را محاسبه می‌کنیم:

$$Av_1 = 1v_1 \quad \Rightarrow \quad v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad , \quad Av_2 = 0 \quad \Rightarrow \quad v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

- بردارویژه‌های v_1 و v_2 مستقل‌اند، زیرا v_1 و v_2 بر هم عمود بوده و ناصفرند.

- ماتریس $S = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$ را در نظر بگیرید. S وارون‌پذیر است و $AS = S\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. بنابراین:

$$S^{-1}AS = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

مثال: فرض کنید $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. آیا A به عنوان ماتریسی در $M_2(R)$ قطری‌شدنی است؟ به عنوان ماتریسی در $M_2(C)$ چطور؟

حل: فرض کنید $S \in M_2(R)$ وجود داشته باشد به طوری که $AS = S\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$. فرض کنید $S = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$. پس :

$$AS = A\begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} Av_1 & Av_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} = \begin{bmatrix} d_1v_1 & d_2v_2 \end{bmatrix}$$

در نتیجه $Av_1 = d_1v_1$ و $Av_2 = d_2v_2$. چون S وارون‌پذیر است پس $v_1, v_2 \neq 0$ بنا به تعریف اعداد حقیقی d_1 و d_2 مقادیر ویژه‌ی A هستند. یعنی ریشه‌های چندجمله‌ای $f(\lambda) = det(\lambda I - A)$. از طرفی چندجمله‌ای $f(\lambda) = det \begin{bmatrix} \lambda & 1 \\ -1 & \lambda \end{bmatrix} = \lambda^2 + 1$ ریشه‌ی حقیقی ندارد. بنابراین فرض وجود $S \in M_2(R)$ باطل است و A به عنوان ماتریسی در $M_2(R)$ قطری‌شدنی نیست.

ریشه‌های چندجمله‌ای $\lambda^2 + 1$ برابر با $\pm i$ است. همچنین اگر $v_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$ و $v_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ را در

نظر بگیرید آنگاه:

$$Av_1 = iv_1 \quad , \quad Av_2 = -iv_2$$

همچنین ماتریس $S = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$ وارون‌پذیر است زیرا $det\ S = -2i \neq 0$. بنابراین $S \in M_2(C)$ وارون‌پذیر وجود دارد که

$$S^{-1}AS = \begin{bmatrix} +i & 0 \\ 0 & -i \end{bmatrix}$$

بنابراین A به عنوان ماتریسی در $M_2(C)$ قطری‌شدنی است. به این معنا که ماتریس وارون‌پذیر $S \in M_2(C)$ وجود دارد به طوری که $S^{-1}AS \in M_2(C)$ و قطری است.

یادداشت: فرض کنید $A \in M_n(F)$ که در آن R یا $F = C$.

قضیه: ماتریس A قطری‌شدنی است اگر و تنها اگر A، n بردار ویژه‌ی مستقل خطی داشته باشد.

برهان:

- $\Leftarrow$: فرض کنید A ماتریس قطری‌شدنی باشد. در این صورت ماتریس $S \in M_n(F)$ وجود دارد به طوری که:

$$S^{-1}AS = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}$$

ستون‌های S را با $v_1, \cdots, v_n$ نمایش می‌دهیم. یعنی $S = \begin{bmatrix} v_1 \cdots v_n \end{bmatrix}$. در این صورت:

$$AS = A\begin{bmatrix} v_1 \cdots v_n \end{bmatrix} = \begin{bmatrix} Av_1 \cdots Av_n \end{bmatrix} = \begin{bmatrix} d_1v_1 \cdots d_nv_n \end{bmatrix}$$

بنابراین به ازای هر $1 \leq i \leq n$، $Av_i = d_iv_i$. چون S وارون‌پذیر است، پس $v_i \neq 0$ لذا d_i و v_i به ازای هر $1 \leq i \leq n$، مقدار ویژه و بردار ویژه‌ی ماتریس A است. چون S وارون‌پذیر است پس $v_1, \cdots, v_n$ مستقل خطی هستند.

- فرض کنید A، n بردار ویژه‌ی مستقل خطی ، $v_1, \cdots, v_n$ ، متناظر با مقادیر ویژه‌ی

λ_n , $\lambda_1, \cdots$ دارد. بنابراین قرار دهید $S = \begin{bmatrix} v_1 \cdots v_n \end{bmatrix}$. در نتیجه:

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

نکته: فرض کنید مقادیر ویژه‌ی ماتریس n در n A متمایز باشند. در این صورت بردار ویژه‌های متناظر مستقل خطی هستند.

برهان خلف: فرض کنید v_k , $v_1, \cdots$ بردار ویژه‌های مستقل باشند به طوری که به ازای هر $k+1 \leq i \leq n$ v_i بردار وابسته به بردارهای v_k , $v_1, \cdots$ باشند به طوری که $k < n$. در این صورت:

$$v_{k+1} = \sum_{j=1}^{k} c_j v_j$$

به طوری که حداقل یکی از c_jها $1 \leq j \leq k$ ناصفر است. با محاسبه‌ی اثر ماتریس A روی بردار v_{k+1} به دست می‌آید که:

$$\lambda_{k+1} v_{k+1} = \sum_{j=1}^{k} c_j \lambda_j v_j \quad (1)$$

که در آن، λ_j مقدار ویژه‌ی متناظر با بردار ویژه‌ی v_j است. از طرفی،

$$\lambda_{k+1} v_{k+1} = \sum_{j=1}^{k} c_j \lambda_{k+1} v_j \quad (2)$$

با کم کردن رابطه‌ی (۱) از (۲)، خواهیم داشت:

$$0 = \sum_{j=1}^{k} c_j (\lambda_{k+1} - \lambda_j) v_j$$

بنا به فرض، حداقل یکی از c_j ها (مثلا c_{j}^{*}) ناصفر است. چون v_1 تا v_k مستقل خطی شدند، $(0 \leq j^* \leq k)$ $\lambda_{k+1} - \lambda_{j^*} = 0$ که در تناقض با متمایز بودن λ_1 تا λ_n است؛ بنابراین، فرض وابسته بودن v_i ها باطل بوده و بردارهای ویژه، مستقل خطی‌اند.

نکته‌ی ۲: همه‌ی ماتریس‌ها، قطری‌شدنی نیستند.

مثال: فرض کنید $A = \begin{bmatrix} ۰ & ۱ \\ ۰ & ۰ \end{bmatrix}$. بنابراین، $f(x) = det(xI - A) = x^۲$ و در نتیجه مقدار ویژه‌ی A برابر $\lambda = ۰$ با تکرار ۲ است. بردار ویژه‌ی متناظر با آن، این‌گونه به دست می‌آید:

$$Ax = ۰ \rightarrow \begin{bmatrix} ۰ & ۱ \\ ۰ & ۰ \end{bmatrix} \begin{bmatrix} x_۱ \\ x_۲ \end{bmatrix} = \begin{bmatrix} x_۲ \\ ۰ \end{bmatrix} \rightarrow x_۲ = ۰$$

بنابراین، فضای ویژه‌ی مربوط به بردار ویژه‌ی آن، یعنی $W = \{v \in R^۲ | v = ۰\}$ فضایی یک‌بعدی است، لذا A دو بردار ویژه‌ی مستقل ندارد و در نتیجه، بنا به قضیه، قطری‌شدنی نیست.

نکته‌ی ۳: اگر ماتریس A قطری‌شدنی باشد، آن‌گاه ماتریس S که $S^{-۱}AS$ قطری است، یکتا نیست.

تعریف: فرض کنید T تبدیل خطی روی فضای σ با بعد متناهی باشد. گوییم T قطری‌شدنی است اگر وجود داشته باشد پایه‌ای مانند B برای V به طوری‌که ماتریس نمایش T در پایه‌ی B که با $[T]_B$ نمایش می‌دهیم، قطری باشد.

یادداشت ۱: فرض کنید T تبدیل خطی قطری‌شدنی باشد و $B = \{v_۱, \cdots, v_n\}$ پایه‌ی مورد نظر باشد، یعنی:

$$[T]_B = \begin{bmatrix} d_۱ & \cdots & ۰ \\ \vdots & \vdots & \vdots \\ ۰ & \cdots & d_n \end{bmatrix}$$

بنابراین، به ازای هر $۱ \leq i \leq n$ داریم $T(v_i) = d_i v_i$، یعنی ستون i ام ماتریس $[T]_B$ مختصات $T(v_i)$ در پایه‌ی B است. در نتیجه، $d_۱$ تا d_n مقادیر ویژه‌ی T بوده و $v_۱$ تا v_n بردارهای ویژه‌ی متناظرشان هستند.

با تعویض عناصر پایه، می‌توان مقادیر ویژه را برحسب تکررشان، مرتب نمود و فرض کرد که $d_۱$ تا d_k برابر k تا مقدار ویژه‌ی مجزا با تکرارهای به ترتیب $n_۱$ تا n_k اند، بنابراین:

$$[T]_{B'} = \begin{bmatrix} \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & d_1 \end{bmatrix} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \begin{bmatrix} d_k & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & d_k \end{bmatrix} \end{bmatrix} = \begin{bmatrix} d_1 I_{n_1} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & d_k I_{nk} \end{bmatrix}$$

در نتیجه، چندجمله‌ای ویژه‌ی تبدیل خطی T برابر است با:

$$f(x) = det(xI - [T]_{B'}) =$$

$$det \begin{bmatrix} xI_{n_1} - d_1 I_{n_1} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & xI_{nk} - d_k I_{nk} \end{bmatrix} = (x - d_1)^{n_1} + \cdots + (x - d_k)^{n_k}$$

به ازای هر i از ۱ تا k فضای ویژه‌ی مربوط به مقدار ویژه‌ی d_i برابر است با $W_i = N(d_i I - T)$ و بنابراین:

$$dimW_i = dimN(\begin{bmatrix} d_i I - d_1 I & \cdots & 0 \\ \vdots & \vdots & \vdots \\ \vdots & d_i I - d_i I = 0 & \vdots \\ \vdots & \vdots & \vdots \\ 0 & \cdots & d_i I - d_k I \end{bmatrix})$$

چون d_i ها مقادیر ویژه‌ی متمایزی هستند، پس بنا به قضیه‌ی رتبه، $dimW_i = n_i$ است.

یادداشت فوق، کم و بیش، قضیه‌ی زیر را به دنبال خواهد داشت:

قضیه: فرض کنید T تبدیل خطی روی فضای با بعد متناهی V است، مقادیر ویژه‌ی متمایز T برابر λ_1 تا λ_k اند و W_i فضای پوچ $T - \lambda_i I$، فضای ویژه‌ی مربوط به مقدار ویژه‌ی λ_i، به ازای $1 \leq i \leq k$ است؛ در این صورت، گزاره‌های زیر معادل‌اند:

- ماتریس T قطری‌شدنی است.
- چندجمله‌ای ویژه‌ی T برابر است با:

$$f(x) = (x-\lambda_۱)^{n_۱} \cdots (x-\lambda_k)^{n_k}$$

که در آن، $dimW_i = n_i$ به ازای i از ۱ تا k.

- $\sum_{i=۱}^{k} dimW_i = dimV$.

لم: فرض کنید T تبدیل خطی روی v است و $Tv = \lambda v$. اگر $f(x)$ یک چندجمله‌ای باشد، آنگاه

$$f(T)v = f(\lambda)v \ .$$

برهان: فرض کنید $f(x) = \sum_{i=۰}^{m} a_i x^i$. از طرفی:

$$T^۲(v) = T(Tv) = T(\lambda v) = \lambda T(v) = \lambda^۲ v$$

$$T^۳(v) = T(T^۲ v) = T(\lambda^۲ v) = \lambda^۳ v$$

$$\vdots$$

$$T^n(v) = T(T^{n-۱} v) = T(\lambda^{n-۱} v) = \lambda^n v$$

در نتیجه:

$$f(T)v = (\sum_{i=۰}^{m} a_i T^i)v = \sum_{i=۰}^{m} a_i T^i(v) = \sum_{i=۰}^{m} a_i \lambda^i v = f(\lambda)v \ .$$

نتیجه: اگر $\lambda_۱, \cdots, \lambda_n$ مقادیر ویژه‌ی ماتریس A باشد، آنگاه $\lambda_۱^k, \cdots, \lambda_n^k$ مقادیر ویژه‌ی ماتریس A^k است.

لم: فرض کنید T تبدیل خطی روی فضای خطی با بعد متناهی V است. اگر $\lambda_۱, \cdots, \lambda_k$ مقادیر ویژه‌ی متمایز T باشند و W_i به ازای هر $۱ \leq i \leq k$، فضای ویژه‌ی مربوط به مقدار ویژه λ_i باشد و $v_۱ + \cdots + v_k = ۰$ آنگاه $v_۱ = \cdots = v_k = ۰$.

برهان: به ازای هر j، $1 \leq j \leq k$ تعریف کنید:

$$f_j(x) = \frac{\prod_{i\neq j, i=1}^{k}(x - c_i)}{\prod_{i\neq j, i=1}^{k}(c_j - c_i)}$$

همچنین چون $v_i \in W_i$، پس $T(v_i) = \lambda_i v_i$. بنا به لم قبل:

$$f_j(T)(v_1 + \cdots + v_k) = f_j(T)v_1 + \cdots + f_j(T)v_k = f_j(\lambda_1)v_1 + \cdots + f_j(\lambda_k)v_k$$

چون $v_1 + \cdots + v_k = 0$ بنابراین:

$$f_j(T)(v_1 + \cdots + v_k) = f_j(T)(0) = 0$$

در نتیجه:

$$f_j(\lambda_1)v_1 + \cdots + f_j(\lambda_k)v_k = 0$$

به وضوح $f_j(\lambda_j) = 1$ و $f_j(\lambda_i) = 0$ به ازای هر $1 \leq i \leq k$ و $i \neq j$. بنابراین $v_j = 0$.

لم: فرض کنید T تبدیل خطی روی فضای با بعد متناهی V است. اگر $\lambda_1, \cdots, \lambda_k$ مقادیرویژه‌ی متمایز T باشند و W_i فضای ویژه‌ی مربوط به مقدار ویژه‌ی λ_i باشد، در این صورت:

$$dim\ (W_1 + \cdots + W_k) = \sum_{i=1}^{k} dim\ W_i$$

برهان: فرض کنید $B_i = \{V_{i1}, \cdots, V_{in_i}\}$ پایه‌ی W_i است که در آن $dim\ W_i = n_i$. ادعا می‌کنیم $\cup_{i=1}^{k} B_i$ پایه‌ای برای زیرفضای $W_1 + \cdots + W_k$ است.

- استقلال عناصر $\cup_{i=1}^{k} B_i$: فرض کنید:

$$(c_{11}v11 + \cdots + c_{1n_1}v_{1n_1}) + \cdots + (c_{k1}vk1 + \cdots + (c_{k1}v_{k1} + \cdots + c_{kn_k}v_{kn_k}) = 0$$

قرار دهید $v_i = c_{i1}v_{i1} + \cdots + c_{in_i}v_{in_i}$. بنابراین $v_1 + \cdots + v_k = 0$ بنا به لم به ازای هر $1 \leq i \leq k$، $v_i = 0$. از طرفی چون B_i پایه‌ای برای W_i است پس $c_{i1} = \cdots = c_{in_i} = 0$.

- فرض کنید $v_۱ + \cdots + v_k \in W_۱ + \cdots + W_k$. در این صورت هر v_i ترکیب خطی از عناصر B_i است. پس $v_۱ + \cdots + v_k$ ترکیب خطی عناصر $\cup_{i=۱}^{k} B_i$ است.
در نتیجه $\cup_{i=۱}^{k} B_i$ پایه‌ای برای $W_۱ + \cdots + W_k$ است و لذا حکم ثابت می‌شود زیرا به ازای هر $i \neq j$، $B_i \cap B_j = \emptyset$:

$$dim\,(W_۱ + \cdots + W_k) = |\cup_{i=۱}^{k} B_i| = \sum_{i=۱}^{k} |B_i| = \sum_{i=۱}^{n} dim\, W_i$$

اثبات قضیه‌ی مربوط به قطرسازی:

- $۱ \Leftarrow ۲$: طی یادداشت جلسه‌ی قبل ثابت شد.

- $۲ \Leftarrow ۳$: می‌دانیم که درجه‌ی چندجمله‌ای برابر با درجه‌ی فضای V است. از طرفی:

$$f(x) = (x - \lambda_۱)^{n_۱} \cdots (x - \lambda_k)^{n_k}$$

بنابراین :

$$dim\, V = n_۱ + \cdots + n_k = dim\, W_۱ + \cdots + dim\, W_k$$

- $۳ \Leftarrow ۱$: $W_۱ + \cdots + W_k \subseteq V$ زیرفضا است. بنا به لم $dim\,(W_۱ + \cdots + W_k) = \sum_{i=۱}^{n} dim\, W_i$. بنا به فرض $\sum_{i=۱}^{n} dim\, W_i = dim\, V$. بنابراین $V = W_۱ + \cdots + W_k$.
اگر B_i پایه‌ای برای W_i باشد، بنا به لم $\cup_{i=۱}^{k} B_i$ پایه‌ای برای $W_۱ + \cdots + W_k = V$ است. بنابراین ماتریس نمایش T در پایه‌ی $B = \cup_{i=۱}^{k} B_i$، $[T]_B$ قطری است.

مثال: فرض کنید $A = \begin{bmatrix} ۵ & -۶ & -۶ \\ -۱ & ۴ & ۲ \\ ۳ & -۶ & -۴ \end{bmatrix}$؛ در این صورت، $A^{۵۵۰}$ را بیابید.

حل:

$$f(x) = det(xI - A) = (x - ۱)(x - ۲)^۲ \rightarrow \lambda_۱ = ۲, \lambda_۲ = ۱$$

$$W_۱ = N(۲I - A) = N(\begin{bmatrix} ۳ & -۶ & -۶ \\ -۱ & ۲ & ۲ \\ ۳ & -۶ & -۷ \end{bmatrix}) \to dimW_۱ = ۲$$

به راحتی می‌توان دید که $v_۱ = \begin{bmatrix} ۲ \\ ۰ \\ ۱ \end{bmatrix}$ و $v_۲ = \begin{bmatrix} ۰ \\ ۱ \\ -۱ \end{bmatrix}$ بردارهای ویژه‌ی متناظر با $\lambda_۱ = ۲$ و $\{v_۱, v_۲\}$ پایه‌ای برای $W_۱$ است.

$$W_۲ = N(I - A) = N(\begin{bmatrix} -۴ & ۶ & ۶ \\ ۱ & -۳ & -۲ \\ -۳ & ۶ & ۵ \end{bmatrix}) \to dimW_۲ = ۱\ ,\ \ v_۳ = \begin{bmatrix} ۳ \\ -۱ \\ ۳ \end{bmatrix}$$

بردار ویژه‌ی $\lambda_۲ = ۱$ است.

بنابراین، چون $dimW_۱ + dimW_۲ = dimV$ پس T قطری‌شدنی است. قرار دهید $S = \begin{bmatrix} v_۱ & v_۲ & v_۳ \end{bmatrix}$ در این صورت، $S^{-۱}AS = \begin{bmatrix} ۲ & ۰ & ۰ \\ ۰ & ۲ & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix}$ بنابراین

$$S^{-۱}A^{۵۵۰}S = \begin{bmatrix} ۲^{۵۵۰} & ۰ & ۰ \\ ۰ & ۲^{۵۵۰} & ۰ \\ ۰ & ۰ & ۱^{۵۵۰} \end{bmatrix} \to A^{۵۵۰} = S \begin{bmatrix} ۲^{۵۵۰} & ۰ & ۰ \\ ۰ & ۲^{۵۵۰} & ۰ \\ ۰ & ۰ & ۱ \end{bmatrix} S^{-۱}$$

یادداشت: فرض کنید ماتریس A قطری‌شدنی باشد. بنا به قضیه، چندجمله‌ای ویژه‌ی A به صورت

$$f(x) = (x - \lambda_۱)^{n_۱} \cdots (x - \lambda_k)^{n_k}$$

است و ماتریس وارون‌پذیر S وجود دارد که

$$S^{-۱}AS = \begin{bmatrix} \lambda_۱ I_{n_۱} & \cdots \\ \vdots & \cdots \\ \cdots & \lambda_k I_{n_k} \end{bmatrix}$$

از طرفی $det(xI - A) = det(xI - S^{-1}AS)$ بنابراین A و $S^{-1}AS$ چندجمله‌ای ویژه‌ی یکسانی دارند، زیرا

$$f(A) = f(S^{-1}AS) = (\begin{bmatrix} \lambda_1 I_{n_1} & \cdots \\ \vdots & \ldots \\ \cdots & \lambda_k I_{n_k} \end{bmatrix} - \lambda_1 I)^{n_1} \cdots (\begin{bmatrix} \lambda_1 I_{n_1} & \cdots \\ \vdots & \ldots \\ \cdots & \lambda_k I_{n_k} \end{bmatrix} - \lambda_k I)^{n_k} = 0$$

حال این سوال مطرح می‌شود که آیا به ازای هر ماتریسی مانند A این خاصیت برقرار است (یعنی $f(A) = 0$ که در آن $f(x)$ چندجمله‌ای ویژه‌ی A است)؟ سوال مشابه به طور واضحی برای هر تبدیل خطی T روی فضای برداری با بعد متناهی n قابل طرح است.

قضیه‌ای توسط کیلی و همیلتون ثابت شده است که نتیجه می‌دهد $f(A) = 0$ (که در آن f چندجمله‌ای ویژه‌ی A است) به ازای هر ماتریسی و مشابها برای هر تبدیل خطی روی فضای برداری با بعد متناهی.

پرسش: فرض کنید $A \in M_n(F)$ و $f(x) = det(xI - A)$ چندجمله‌ای ویژه‌ی A باشد. اشکال اثبات زیر را بیابید:

$$f(A) = det(AI - A) = det(0) = 0$$

برای اثبات قضیه‌ی کیلی ـ همیلتون، باید چندجمله‌ای $f(x) = det(xI - A)$ را باز کنیم.

قضیه‌ی کیلی ـ همیلتون:

اگر $A \in M_n(F)$ و F برابر R یا Φ بوده و $f(x)$ چندجمله‌ای ویژه‌ی A باشد، آن‌گاه $f(A) = 0$.

برهان:

فرض کنید

$$f(x) == det(xI - A) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

از طرفی، می‌دانیم که

$$(xI - A)adj(xI - A) = det(xI - A)I \quad (*)$$

بنابراین به محاسبه‌ی درایه‌های $adj(xI - A)$ می‌پردازیم. برای محاسبه‌ی درایه‌ی ij ام $adj(xI - A)$ باید سطر j ام و ستون i ام ماتریس $(xI - A)$ را حذف کنیم، سپس دترمینان آن را محاسبه کرده و آن‌گاه در $(-۱)^{i+j}$ ضرب کنیم. در نتیجه، درایه‌ی ij ام $adj(xI - A)$ یک چندجمله‌ای از درجه‌ی حداکثر $n - ۱$ بر حسب x است، پس می‌توان ماتریس $adj(xI - A)$ را برحسب جملات ۱ و x و... و $x^{n-۱}$ نوشت، یعنی:

$$adj(xI - A) = B_{n-۱}x^{n-۱} + B_{n-۲}x^{n-۲} + \cdots + B_۱x + B_۰$$

به‌طوری‌که $B_i \in M_n(F)$ به ازای هر i که $۰ \leq i \leq n - ۱$ بنا به رابطه‌ی (*) داریم:

$$(xI-A)(B_{n-۱}x^{n-۱}+B_{n-۲}x^{n-۲}+\cdots+B_۱x+B_۰) = (x^n+a_{n-۱}x^{n-۱}+\cdots+a_۱x+a_۰)I$$

از اتحاد فوق استفاده می‌کنیم و روابط زیر استخراج می‌شود:

$$B_{n-۱} = I \rightarrow A^nB_{n-۱} = A^n$$

$$B_{n-۲} - AB_{n-۱} = a_{n-۱}I \rightarrow A^{n-۱}B_{n-۲} - A^nB_{n-۱} = a_{n-۱}A^{n-۱}$$

$$B_{n-۳} - AB_{n-۲} = a_{n-۲}I \rightarrow A^{n-۲}B_{n-۳} - A^{n-۱}B_{n-۲} = a_{n-۲}A^{n-۲}$$

$$\vdots$$

$$B_۰ - AB_۱ = a_۱I \rightarrow AB_۰ - A^۲B_۱ = a_۱A$$

$$-AB_۰ = a_۰I \rightarrow -AB_۰ = a_۰I$$

و در نتیجه، با جمع طرف راست روابط فوق، داریم:

$$۰ = A^n + a_{n-۱}A^{n-۱} + \cdots + a_۰I = f(A)\ .$$

فرض کنید $A \in M_n(F)$، C یا $F = R$. در این صورت $f(A) = ۰$ که در آن $f(x)$ چندجمله‌ای ویژه ماتریس A است.

برای محاسبه‌ی A^k، $k > n$ با استفاده از الگوریتم تقسیم داریم که:

$$x^k = q(x)f(x) + r(x) \qquad deg\, r(x) < n \quad \text{یا} \quad r(x) = ۰$$

بنابراین:

$$A^k = q(A)f(A) + r(A) = r(A)$$

در نتیجه $A^k = r(A)$ که در آن $deg\, r(x) =< n$. به عبارتی، توان kام ماتریس A را می‌توان به صورت ترکیب خطی A_i ها که $i < n$ نوشت.

بنابراین این سوال مطرح می‌شود که آیا چندجمله‌ای با درجه‌ی کمتر از n وجود دارد که $p(A) = ۰$.

تعریف: فرض کنید $A \in M_n(F)$ که F برابر R یا C است؛ در این صورت، چندجمله‌ای ناصفر $g(x)$ که $g(A) = ۰$ را چندجمله‌ای پوچساز A گویند. چندجمله‌ای پوچساز $p(x)$ را که کم‌ترین درجه را دارد و تکین است، چندجمله‌ای مینیمال گویند.

نکته: فرض کنید V فضای خطی با بعد متناهی و T تبدیل خطی روی V است. B و B' را دو پایه‌ی مختلف برای V در نظر بگیرید. در این صورت، ماتریس وارون‌پذیر P وجود دارد که $[T]_B = P[T]_{B'}P$. نشان می‌دهیم که چندجمله‌ای مینیمال $[T]_B$ و $[T]_{B'}$ یکسان است. فرض کنید $p(x) = \sum_{i=۰}^{m} c_i x^i$ چندجمله‌ای مینیمال $[T]_B$ است. آن‌گاه:

$$p(A) = \sum_{i=۰}^{m} c_i A^i = ۰ \quad A = [T]_B$$

$$P^{-۱}p(A)P = P^{-۱}(\sum_{i=۰}^{m} c_i A^i)P = \sum_{i=۰}^{m} c_i P^{-۱}A^iP = \sum_{i=۰}^{m} c_i (P^{-۱}AP)^i = \sum_{i=۰}^{m} c_i [T]_{B'}^i = ۰$$

تعریف: فرض کنید V فضای خطی روی F با بعد متناهی و $T : V \to V$ یک تبدیل خطی است. در این صورت هر چندجمله‌ای ناصفر مانند $g(x)$ که $g(T) = ۰$ را چندجمله‌ای پوچساز گویند.

تعریف: فرض کنید V فضای خطی روی F با بعد متناهی و $T : V \to V$ یک تبدیل خطی است. در این صورت چندجمله‌ای $p(x)$ با ضرایب F را چندجمله‌ای مینیمال گویند هر گاه $p(x)$ چندجمله‌ای تکین باشد (ضریب جمله با بزرگترین درجه آن یک باشد) و همچنین در میان پوچسازهای T، کم‌ترین درجه را داشته باشد.

گزاره: فرض کنید V فضای خطی روی F با بعد متناهی و $T : V \to V$ یک تبدیل خطی است. آن‌گاه، چندجمله‌ای مینیمال T یکتاست.

برهان: فرض کنید $p_۱(T) = ۰$ و $p_۲(T) = ۰$ که در آن

$$p_1(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_0$$

$$p_2(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_0$$

$$\to p_1(x) - p_2(x) = (a_{m-1} - b_{m-1})x^{m-1} + \cdots + (a_0 - b_0)$$

$$(p_1 - p_2)(T) = 0\ ,\quad deg(p_1 - p_2) \leq m - 1$$

توجه کنید که چون p_1 و p_2 هردو چندجمله‌ای مینیمال فرض شده‌اند، پس باید درجه‌ی یکسانی داشته باشند.

در نتیجه، $p_1 - p_2 = 0$.

گزاره: فرض کنید V فضای خطی روی F با بعد متناهی، $T : V \to V$ یک تبدیل خطی و $g(x)$ یک چندجمله‌ای پوچساز است؛ در این صورت، $p(x)|g(x)$ که در آن $p(x)$ چندجمله‌ای مینیمال است.

برهان: با استفاده از الگوریتم تقسیم، چندجمله‌ای $q(x)$ و $r(x)$ وجود دارد به طوری که $g(x) = q(x)p(x) + r(x)$ که $degr(x) < degp(x)$ یا $r(x) = 0$.

توجه کنید که چون $p(x)$ چندجمله‌ای مینیمال است، پس $degp(x) \leq degg(x)$ و در نتیجه $g(T) = q(T)p(T) + r(T) = r(T) = 0$

اگر $degr(x) < degp(x)$ به تناقض با مینیمال بودن $p(x)$ می‌رسیم، پس $r(x) = 0$ و در نتیجه:

$$g(x) = q(x)p(x) \to p(x)|g(x)\ .$$

یادآوری: فرض کنید $M \in M_n(F)$ و C یا $F = R$. در این صورت $p(x)|f(x)$ که در آن $p(x)$ و $f(x)$ به ترتیب چندجمله‌ای مینیمال و چندجمله‌ای ویژه هستند. به طور مشابه به ازای هر تبدیل خطی روی فضای خطی با بعد متناهی، این گزاره برقرار است.

قضیه: فرض کنید T تبدیل خطی روی فضای خطی V با بعد متناهی n است. در این صورت چندجمله‌ای مینیمال و چندجمله‌ای ویژه T (و یا هر ماتریس A) دارای ریشه‌های یکسان هستند ولی احتمالاً چندگانگی‌های متفاوت دارند.

برهان: فرض کنید λ ریشه‌ی چندجمله‌ای ویژه‌ی T باشد. در این صورت λ مقدارویژه‌ی T است. لذا وجود دارد بردار ناصفر $V \ni x \neq 0$ به طوری که $Tx = \lambda x$. بنابراین $p(T)x = p(\lambda)x$. چون $p(x)$ چندجمله‌ای مینیمال است، لذا $p(\lambda)x = 0$. چون $x \neq 0$ در نتیجه $p(\lambda) = 0$. پس λ ریشه‌ی $p(x)$ است.

حال فرض کنید λ ریشه‌ی $p(x)$ است. در نتیجه $p(\lambda) = ۰$. بنابراین $p(x) = (x-\lambda)q(x)$ که در آن $q(x)$ چندجمله‌ای با ضرایب F است. بنابراین:

$$۰ = p(T) = (T - \lambda I)q(T)$$

بنابراین $(T-\lambda I)q(T) = ۰$. از طرفی $deg\ (q(x)) < deg\ (p(x))$. بنابراین چون $p(x)$ چندجمله‌ای مینیمال T است، نتیجه می‌گیریم که $q(T) \neq ۰$. پس وجود دارد بردار ناصفر $x \in V$ که $q(T)x \neq ۰$. از طرفی $(T-\lambda I)q(T)x = ۰$. قرار دهید $y = q(T)x$، لذا $y \neq ۰$ و $(T-\lambda I)y = ۰$. در نتیجه $Ty = \lambda y$ و $y \neq ۰$. لذا λ مقدارویژه‌ی T است و لذا λ ریشه‌ی چندجمله‌ای ویژه‌ی T است.

نکته: فرض کنید $A \in M_n(C)$ بنابراین $f(x)$ و $p(x)$ چندجمله‌ای‌هایی با ضرایب حقیقی هستند. چندجمله‌ای $f(x)$ از درجه n، در اعداد مختلط (با احتساب تکرر) n ریشه دارد. فرض کنید $\lambda_۱, \cdots, \lambda_k$ ریشه‌های حقیقی و $\lambda_{k+۱}, \cdots, \lambda_n$ ریشه‌های مختلط آن باشند. در این صورت:

$$f(x) = (x-\lambda_۱)\cdots(x-\lambda_k)(x-\lambda_{k+۱})\cdots(x-\lambda_n)$$

حال ادعا می‌کنیم که اگر λ_i، $k+۱ \leq i \leq n$، ریشه‌ی مختلط $f(x)$ باشد، آنگاه $\bar{\lambda_i}$ نیز ریشه $f(x)$ است. فرض کنید:

$$f(x) = x^n + a_{n-۱}x^{n-۱} + \cdots + a_۱x + a_۰ \quad , f(\lambda_i) = ۰, a_i \in R$$

لذا

$$f(\lambda_i) = \lambda_i^n + a_{n-۱}\lambda^{n-۱} + \cdots + a_۱\lambda_i + a_۰ = ۰$$

از طرفین تساوی فوق مزدوج می‌گیریم. چون $a_i \in R$، به ازای $۰ \leq i \leq n-۱$ داریم:

$$f(\bar{\lambda_i}^n + a_{n-۱}\bar{\lambda_i}^n - ۱ + \cdots + a_۱\bar{\lambda_i} + a_۰ = ۰$$

بنابراین $\bar{\lambda_i}$ نیز ریشه‌ی $f(x)$ است. لذا $(x-\lambda_i)(x-\bar{\lambda_i})|f(x)$ و

$$(x-\lambda_i)(x-\bar{\lambda_i}) = x^۲ - (\lambda_i + \bar{\lambda_i})x + \lambda_i\bar{\lambda_i} = x^۲ + a'_ix + b'_i \quad a'_i, b'_i \in R$$

بنابراین $f(x)$ را می‌توان به صورت حاصلضرب تعدادی عوامل درجه یک، $x - \lambda_i$، به ازای $۱ \leq$

$i \leq k$ و تعدادی عوامل درجه دوم نوشت.

$$f(x) = (x - \lambda_1) \cdots (x - \lambda_k)(x^2 + a'_i x + b'_i) \cdots (x^2 - a'_m x + b'_m)$$

که $m = \frac{n-k}{2}$. عوامل درجه دوم روی R تحویل ناپذیرند، یعنی نمی‌توان آن‌ها را به عوامل درجه اول تجزیه کرد. از طرفی هر یک از عوامل $(x - \lambda_i)$ و $x^2 + a'_j x + b'_j$ ممکن است در $f(x)$ تکرر داشته باشند. مستقل از تکرر، هر عامل $f(x)$ عامل $p(x)$ است و هر عامل $p(x)$، عامل $f(x)$ است.

نتیجه: اگر T تبدیلی خطی قطری‌شدنی باشد و $\lambda_1, \cdots, \lambda_l$ مقادیر ویژه متمایز T باشند، آنگاه:

$$p(x) = (x - \lambda_1) \cdots (x - \lambda_l)$$

یعنی چندجمله‌ای مینیمال یک تبدیل خطی قطری‌شدنی ریشه‌ی تکراری ندارد.

برهان: فرض کنید W_i فضای ویژه‌ی مربوط به مقدار ویژه‌ی λ_i $(1 \leq i \leq k)$ باشد، آنگاه چون T قطری‌شدنی است بنا به قضیه

$$V = W_1 + \cdots + W_k$$

فرض کنید $x_j \in W_j$. بنابراین $(T - \lambda_j I)x_j = 0$. در نتیجه چون T تبدیل قطری‌شدنی است، پس

$$(T - \lambda_1 I) \cdots (T - \lambda_k I)x_j = (T - \lambda_1 I) \cdots (T - \lambda_j I)x_j = 0 \quad 1 \leq j \leq k$$

از طرفی اگر $x \in V$ آنگاه $x = x_1 + \cdots + x_k$ که در آن $x_j \in W_j$. در نتیجه

$$\prod_{j=1}^{k}(T-\lambda_j I)x = \prod_{j=1}^{k}(T-c_j I)(x_1+\cdots+x_k) = \prod_{j=1}^{k}(T-c_j I)x_1+\cdots+\prod_{j=1}^{k}(T-c_j I)x_k$$

$$= 0 + \cdots + 0 = 0$$

بنابراین $\prod_{j=1}^{k}(T - \lambda_j I) = 0$ در نتیجه $p(x) | \prod_{j=1}^{k}(x - \lambda_j I)$. از طرفی بنا به قضیه ریشه‌های چندجمله‌ای مینیمال و چندجمله‌ای ویژه یکسان هستند (مگر و احتمالاً در تکرر)؛ پس $\prod_{j=1}^{k}(x - \lambda_j) | f(x)$. از طرفی چون $p(x)$ تکین است، پس $p(x) = \prod_{j=1}^{k}(x - \lambda_j)$.

ماتریس (تبدیل خطی) مثلثی شدنی

تعریف: ماتریس $A \in M_n(F)$ که $F = R, C$ را مثلثی‌شدنی گویند هر گاه ماتریس وارون‌پذیر S وجود داشته باشد به طوری که $S^{-۱}AS$ ماتریسی بالامثلثی یا پایین‌مثلثی باشد. تبدیل خطی T روی فضای V با بعد متناهی n را مثلثی‌شدنی گویند هرگاه وجود داشته باشد پایه‌ای برای V مانند B به طوری که $[T]_B$ ماتریس بالامثلثی یا پایین‌مثلثی باشد.

قضیه: فرض کنید T تبدیل خطی روی V با بعد متناهی n است؛ در این صورت، T مثلثی‌شدنی است اگر و تنها اگر چندجمله‌ای مینیمال T به چندجمله‌ای‌های از درجه‌ی یک تجزیه شود.

برهان: فرض کنید T مثلثی‌شدنی باشد، پس پایه‌ی B برای T وجود دارد به طوری که $[T]_B$ ماتریس بالامثلثی است:

$$[T]_B = \begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ ۰ & a_{۱۲} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ ۰ & \cdots & \cdots & a_{nn} \end{bmatrix}$$

$$f(x) = det(xI - [T]_B) = det(xI - \begin{bmatrix} a_{۱۱} & a_{۱۲} & \cdots & a_{۱n} \\ ۰ & a_{۱۲} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ ۰ & \cdots & \cdots & a_{nn} \end{bmatrix})$$

$$= det \begin{bmatrix} x - a_{۱۱} & -a_{۱۲} & \cdots & -a_{۱n} \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & x - a_{nn} \end{bmatrix} = (x - a_{۱۱}) \cdots (x - a_{nn})$$

چون چندجمله‌ای مینیمال $p(x)$ بر $f(x)$ قابل تقسیم است، پس چندجمله‌ای مینیمال T به چندجمله‌ای‌های از درجه‌ی یک تجزیه می‌شود.

جهت دیگر برهان:

با استقرا روی $dimV = n$ حکم را ثابت می‌کنیم.

اگر $dimV = ۱$، بدیهی است؛ پس فرض کنید $dimV > ۱$ و

$$p(x) = (x - \lambda_۱)^{r_۱} \cdots (x - \lambda_k)^{r_k}$$

که در آن $\lambda_i \in F$. چون $V \neq \{۰\}$ پس وجود دارد $۰ \neq y \in V$. به وضوح $p(T)y = ۰$. فرض کنید $g(x)$ چندجمله‌ای با ضرایب F است با کم‌ترین درجه‌ی تکین، به طوری‌که $g(T)y = ۰$. ادعا می‌کنیم که $g(x)|p(x)$ زیرا

$$p(x) = q(x)g(x) + r(x)$$

$$s.t \quad degr(x) < degg(x) \;\; or \;\; r(x) = ۰$$

از طرفی $p(T) = g(T)q(T) + r(T)$ و $۰ = p(T)y = q(T)y(T)y + r(t)y$ و $g(T)y = ۰$ بنابراین $r(T)y = ۰$ اگر $degr(x) < degg(x)$ و $r(x) \neq ۰$ تناقض است با انتخاب $g(x)$، پس $g(x)|p(x)$ است. چون $g(x)$ تکین انتخاب شده، پس $g(x) \neq ۰$ لذا $g(x) = (x - \lambda_j)hh(x)$ و هم‌چنین

$$g(T)y = (T - \lambda_j I)h(T)y$$

قرار دهید $h(T)y = w$ و لذا $(T - \lambda_j I)w = ۰$، بنابراین w بردار ویژه‌ی T متناظر با مقدار ویژه‌ی λ_j است و در نتیجه $Tw = \lambda_j w$. حال w را به یک پایه مانند $B = \{w_۱, v_۲, \cdots, v_n\}$ برای V گسترش می‌دهیم؛ در این صورت،

$$[T]_B = \begin{bmatrix} \lambda_j & * & \cdots & * \\ ۰ & & & \\ ۰ & & A & \\ ۰ & & & \end{bmatrix}$$

که در آن A ماتریسی $(n - ۱) \times (n - ۱)$ است. از طرفی، با ضرب بلوکی ماتریس‌ها به ازای هر چندجمله‌ای با ضرایب F،

$$g([T]_B) = \begin{bmatrix} g(\lambda_j) & * & \cdots & * \\ ۰ & & & \\ ۰ & & g(A) & \\ ۰ & & & \end{bmatrix}$$

بنابراین اگر $p(x)$ چندجمله‌ای مینیمال T باشد، آنگاه:

$$۰ = p([T]_B) = \begin{bmatrix} p(\lambda_j) & * & \cdots & * \\ ۰ & & & \\ ۰ & & p(A) & \\ ۰ & & & \end{bmatrix}$$

بنابراین $p(A) = ۰$. در نتیجه، چندجمله‌ای مینیمال A، چندجمله‌ای مینیمال T را می‌شمارد و لذا چندجمله‌ای مینیمال A نیز به عوامل درجه‌ی یک قابل تجزیه است، چون A ماتریسی $(n - ۱) \times (n - ۱)$ است. طبق فرض استقرا، وجود دارد پایه‌ای که در آن پایه‌ی A ماتریس بالامثلثی است (توجه کنید که S وارون‌پذیر است و $S^{-۱}AS$ مثلثی شدنی است و ستون‌های S تشکیل پایه می‌دهند)، یعنی:

$$S^{-۱}AS = \begin{bmatrix} \lambda_{i_۱} & & & * \\ & \lambda_{i_۲} & & \\ \vdots & \vdots & \vdots & \vdots \\ ۰ & & & \lambda_{i_n} \end{bmatrix}$$

قرار دهید

$$Q = \begin{bmatrix} ۱ & ۰ & \cdots & ۰ \\ ۰ & & & \\ ۰ & & P & \\ ۰ & & & \end{bmatrix}$$

در این صورت،

$$Q^{-۱} = \begin{bmatrix} ۱ & ۰ & \cdots & ۰ \\ ۰ & & & \\ ۰ & & P^{-۱} & \\ ۰ & & & \end{bmatrix}$$

و هم‌چنین،

$$Q^{-1}[T]_BQ = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ 0 & & P^{-1} & \\ 0 & & & \end{bmatrix} \begin{bmatrix} \lambda_j & * & \cdots & * \\ 0 & & & \\ 0 & & A & \\ 0 & & & \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ 0 & & P & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} \lambda_j & * & \cdots & * \\ 0 & \lambda_{i_1} & \cdots & * \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{i_n} \end{bmatrix}$$

نکته: اگر ماتریسی، بالامثلثی باشد، حتما یک پایه وجود دارد که در آن، پایین مثلثی است؛ کافی است ترتیب پایه‌ها را کاملا عوض کنیم.